



UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO

Centro de Tecnologia e Ciências

Faculdade de Engenharia

Evelyn da Silva Viana

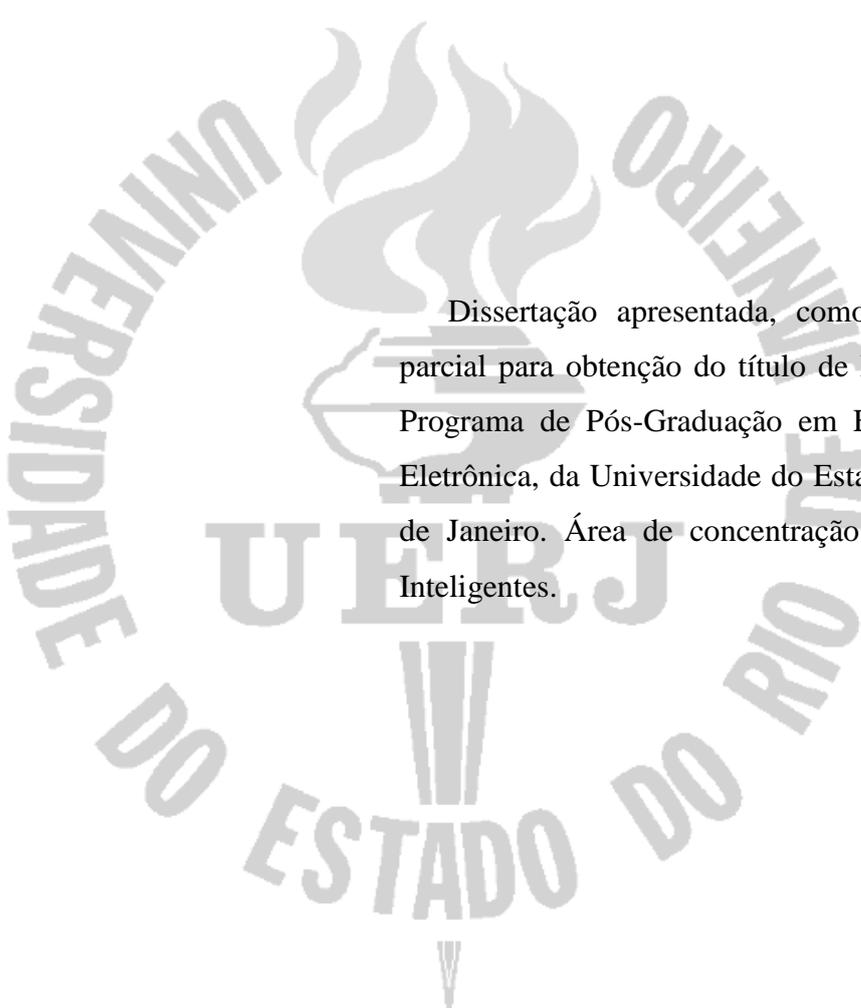
Estimação de Velocidade de Composições Usando Processamento de Vídeo

Rio de Janeiro

2011

Evelyn da Silva Viana

Estimação de Velocidade de Composições Usando Processamento de Vídeo



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Sistemas Inteligentes.

Orientador: Prof. D.Sc. Lisandro Lovisolo

Rio de Janeiro

2011

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/CTC/B

AAA Viana, Evelyn da Silva.
Estimação de Velocidade de Composições Usando Processamento de
Vídeo /
Evelyn da Silva Viana. - 2011.
XXf.

Orientador: Prof. Dr. Lisandro Lovisolo
Dissertação (Mestrado) - Universidade do Estado do Rio de Janeiro,
Faculdade de Engenharia.
Bibliografia:

1.

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação.

Assinatura

Data

Evelyn da Silva Viana

Estimação de Velocidade de Composições Usando Processamento de Vídeo

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Sistemas Inteligentes.

Aprovado em: _____

Banca Examinadora:

Prof. D.Sc. Lisandro Lovisolo (Orientador)
Faculdade de Engenharia da UERJ

Prof.^a D.Sc. Maria Luiza Fernandes Velloso
Faculdade de Engenharia da UERJ

Prof.^a Dr.^a Karla Tereza Figueiredo Leite
Pontifícia Universidade Católica do Rio de Janeiro

Prof. Dr. Alexandre Gomes Ciancio
USC

Rio de Janeiro
2011

DEDICATÓRIA

À minha família, com a qual eu aprendo valores como união, força, perseverança, amor e fé.

AGRADECIMENTOS

À Deus, que carinhosamente esteve comigo me dando ânimo, força e mostrando os caminhos a serem trilhados ao longo não só desta jornada, mas de toda a minha vida.

Agradeço aos meus pais, Jorge e Sandra, pelo amor, incentivo e apoio ao longo de toda minha vida, sempre acreditando no meu potencial. Por todo o carinho e dedicação, principalmente, nos momentos em que o desânimo e o cansaço pareciam grandes demais. Ao meu irmão Mauricio, que do seu jeito sempre me incentivou e apoiou.

Ao professor orientador, Lisandro Lovisolo, por não me deixar desistir, por todas as suas idéias, por sua disposição e confiança.

À Flavio Considera, grande amigo por quem eu tenho grande carinho, admiração e respeito, por seu inesgotável apoio e disposição, por todas as idéias e a quem devo a conclusão deste trabalho.

A Cynthia Kisse por sua amizade ao longo desses anos de luta, sua presença constante nos momentos de alegrias e tristezas, por ser uma fonte de apoio e incentivo constante, sempre solícita e inspiradora.

À Alexandre Custódio, principalmente por sua amizade, por todo o seu apoio e ensinamentos, servindo como inspiração ao longo da caminhada.

Aos queridos companheiros de trabalho em especial a Ricardo Custódio, Délio Cunha e Arnaldo Constâncio por sua paciência e boa vontade contribuindo para a conclusão deste projeto.

Aos amigos, que, de alguma forma me ajudaram nesta jornada e todos aqueles que mesmo sem conhecer se dispuseram a ajudar.

Enfim, a todos aqueles que apoiaram minha empreitada no mestrado e que agora podem se alegrar comigo por essa conquista.

Há mais pessoas que desistem do que pessoas que fracassam.

Henry Ford

RESUMO

Viana, Evelyn da Silva. **Estimação de Velocidade de Composições Usando Processamento de Vídeo**. 2011. 60f. Dissertação (Mestrado em Engenharia Eletrônica) - Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2011.

Esta dissertação tem como objetivo estudar um método para estimar a velocidade de veículos ferroviários usando processamento de vídeos. O sistema proposto consiste na utilização de câmeras ao longo das vias férreas que permitam não só o monitoramento do tráfego ferroviário, mas cujo vídeo captado pode ser utilizado para a obtenção de estimativas para a velocidade instantânea dos trens que por ela trafegam. Tal sistema seria utilizado independentemente dos sistemas de controle já utilizados pela operadora do sistema ferroviário, permitindo que os controladores possam ter uma segunda análise no caso de falha da primeira, assim como um modelo que permita avaliar a velocidade instantânea do veículo ferroviário ao longo do percurso. Os algoritmos de rastreamento empregados para esse fim abordaram diferentes métodos. Confrontou-se os resultados obtidos com dados empíricos de forma a determinar aquele com melhor resposta dadas as características do sistema. O algoritmo que apresentou os melhores resultados emprega um único bloco de referência, mantido em uma posição fixa ao longo de todos os quadros comparados. A métrica de similaridade responsável por determinar quais blocos são mais ou menos similares dentro do universo de busca estipulado é a soma de diferenças absolutas (*SAD, Sum of Absolute Differences*). O tempo de processamento requerido por cada um dos métodos de rastreamento estudados também foi considerado nas avaliações de resultados apresentadas. Uma comparação realizada entre as velocidades coletadas e aquelas informadas pelo sistema de controle mostraram que os resultados obtidos com o sistema atual, utilizando a sinalização apenas por circuito de via apresenta resultados pouco confiáveis com erros bastante significativos. Os resultados obtidos com o sistema proposto apresentaram erros menores quando comparados àqueles obtidos pelo sistema vigente, apresentando-se assim como uma solução viável e de baixo custo quando comparada às técnicas atualmente empregadas para a medida de velocidade de trens.

Palavras-chave: Detecção de Movimento, SAD, Velocidade do Trem

ABSTRACT

This dissertation aims to study a method to estimate the speed of railway vehicles using video processing. The use of cameras along the railways not only enabling the monitoring of traffic, but whose video captured pose be used to obtain estimates of the speed of trains traveling through it. Such system would be used regardless of the control systems already used by the operator of the rail system, allowing the drivers may have a second analysis in case of failure of the first, as a model to assess the speed of rail vehicle along the route .

Different methods of tracking algorithms were employed for this purpose. The results were confronted with empirical data to determine the one with the best answer given. The algorithm presented the best results were the one that employs a single reference block, kept in a fixed position over all quadros compared. The similarity metric responsible for determining which blocks are more or less similar within the universe of search set is the sum of absolute differences.

The results obtained with the proposed system had errors smaller compared to those obtained by the current system, presenting itself as a viable and cost-effective when compared to the techniques currently employed to measure the speed of trains.

Keywords: Motion Detection, SAD, Train Velocity

LISTA DE FIGURAS

Figura 1. Diagrama de blocos da utilização do sistema	17
Figura 2. Sinalização Fixa com sinal simples.....	18
Figura 3. Sinalização Fixa com sinais triplos e repetidores	19
Figura 4. Circuito de Via	19
Figura 5. Visão do Sistema Proposto	29
Figura 6. Distâncias entre o dispositivo e os objetos.....	31
Figura 7. Captura da Imagem	32
Figura 8. Formação de uma imagem	33
Figura 9. Detecção do Movimento	34
Figura 10. Comparação entre quadros	36
Figura 11. Padrão de uma janela identificada.....	39
Figura 12. Padrões dos Trens	39
Figura 13. Sistema com bloco de referência estático	40
Figura 14. Sistema com bloco de referência dinâmico.....	41
Figura 15. Comparação entre blocos	43
Figura 16. Velocidade dos blocos estáticos ao longo do vídeo	57
Figura 17. Gráfico dos Erros obtidos para bloco único dinâmico de tamanho 48x48	60
Figura 18. Gráfico dos erros obtidos para bloco único dinâmico de tamanho 64x64	61
Figura 19. Gráfico dos erros com único bloco para os diferentes tipos de algoritmo	67
Figura 20. Gráfico dos erros com dois blocos para os diferentes tipos de algoritmo.....	68

LISTA DE TABELAS

Tabela 1. Levantamento dos trens	30
Tabela 2. Resultados obtidos com o filtro Gaussiano	47
Tabela 3. Resultados obtidos com a máscara de nitidez e Filtagem <i>High-Boost</i>	47
Tabela 4. Resultados obtidos com o Filtro de Medianas	48
Tabela 5. Tempo de processamento obtidos para o bloco único estático com diferentes tamanhos	49
Tabela 6. Resultados médios obtidos para o Bloco único Estático de 8x8	51
Tabela 7. Resultados médios obtidos para o bloco único estático de 16x16.....	52
Tabela 8. Resultados médios obtidos para o Bloco único Estático de 32x32	53
Tabela 9. Resultados médios obtidos para o bloco único estático de tamanho 48x48.....	53
Tabela 10. Resultados médios obtidos para o bloco único estático de tamanho 64x64.....	54
Tabela 11. Resultados médios obtidos para o bloco único estático de tamanho128x128.....	55
Tabela 12. Tempos de processamento obtidos para dois blocos estáticos com diferentes tamanhos	56
Tabela 13. Resultados médios obtidos para dois blocos estáticos de tamanho 48x48.....	57
Tabela 14. Resultados médios obtidos para dois blocos estáticos de tamanho 64x64.....	58
Tabela 15. Tempos de processamento obtidos para o bloco único móvel com diferentes tamanhos	59
Tabela 16. Resultados médios obtidos para o bloco único móvel de tamanho 48x48	60
Tabela 17. Resultados médios obtidos para o bloco único estático de tamanho 64x64.....	61
Tabela 18. Tempos de processamento obtidos para dois blocos móveis com diferentes tamanhos	62
Tabela 19. Resultados médios obtidos para dois blocos móveis de tamanho 48x48	63
Tabela 20. Resultados médios obtidos para dois blocos móveis de tamanho 64x64	63

Tabela 21. Tempos de processamento obtidos para o bloco médio único com diferentes tamanhos	64
Tabela 22. Resultados médios obtidos para o bloco médio único de tamanho 48x48	64
Tabela 23. Resultados médios obtidos para bloco médio único de tamanho 64x64	65
Tabela 24. Tempos de processamento obtidos para dois blocos médios com diferentes tamanhos	66
Tabela 25. Resultados médios obtidos para dois blocos médios de tamanho 48x48	66
Tabela 26. Resultados médios obtidos para dois blocos médios de tamanho 64x64	67
Tabela 27. Erros obtidos para os trens em diferentes direções e com inclinação.....	69
Tabela 28. Erro obtido para diferentes distâncias.....	69
Tabela 29. Erro na estimação da velocidade para os sistemas analisados.....	70

SUMÁRIO

	INTRODUÇÃO	14
	Motivação	16
	Objetivo	17
1	SISTEMA DE SINALIZAÇÃO FERROVIÁRIO	18
2	ANÁLISE E PROCESSAMENTO DOS VÍDEOS	21
2.1	Movimento	22
2.2	Detecção de Movimento	23
2.3	Redução de Ruído	25
2.4	Métrica de Similaridade	27
3	SISTEMA PROPOSTO	29
3.1	Características do Sistema	29
3.2	O Banco de Dados	30
3.3	Parâmetros de Captura dos Trens	31
3.4	Imagem de Referência	34
3.5	Tamanho da Janela	36
3.6	Tamanho e posicionamento dos blocos	37
3.7	O Rastreamento	38
3.7.1	<u>Bloco Estático</u>	40
3.7.2	<u>Bloco Dinâmico</u>	41
3.8	Comparação entre Blocos	42
3.9	Determinação da Velocidade	43
3.9.1	<u>Validação dos Testes</u>	44
3.9.2	<u>Estimativa de Erro</u>	44
4	RESULTADOS	46
4.1	Comparação dos resultados obtidos para os filtros	46
4.2	Bloco único estático	49
4.2.1	<u>Bloco de 8x8 pixels</u>	49
4.2.2	<u>Bloco de 16x16 pixels</u>	51
4.2.3	<u>Bloco de 32x32 pixels</u>	52
4.2.4	<u>Bloco de 48x48 pixels</u>	53
4.2.5	<u>Bloco de 64x64 pixels</u>	54
4.2.6	<u>Bloco de 128x128 pixels</u>	54
4.3	Dois Blocos Estáticos	55
4.3.1	<u>Bloco de 48x48 pixels</u>	56
4.3.2	<u>Bloco de 64x64 pixels</u>	57
4.4	Bloco Único Dinâmico	58
4.4.1	<u>Bloco de 48x48 pixels</u>	59
4.4.2	<u>Bloco de 64x64 pixels</u>	60
4.5	Dois Blocos Móveis	61
4.5.1	<u>Bloco de 48x48 pixels</u>	62
4.5.2	<u>Bloco de 64x64 pixels</u>	63

4.6	Bloco Médio Único	63
4.6.1	<u>Bloco de 48x48 pixels</u>	64
4.6.2	<u>Bloco de 64x64 pixels</u>	65
4.7	Dois Blocos Médios	65
4.7.1	<u>Bloco de 48x48 pixels</u>	66
4.7.2	<u>Bloco de 64x64 pixels</u>	67
4.8	Comparação dos resultados obtidos para diferentes direções e distâncias	68
	CONCLUSÃO	71

INTRODUÇÃO

Desde seu início a humanidade sentiu a necessidade de deslocamento. Com a mudança das características da sociedade e o desenvolvimento do comércio surgiu a necessidade de uma maneira de interligar diferentes regiões. Este fato permitiu a evolução dos meios de transporte utilizando tração animal até os diferentes meios de transportes conhecidos atualmente.

Os principais meios de transporte utilizados hoje são o aeroviário, rodoviário, ferroviário, aquaviário, tais formas tornaram-se indispensáveis à vida moderna devido a sua agilidade e comodidade de seus passageiros, Todos estes métodos de locomoção, porém trouxeram consigo problemas, até então, pouco conhecidos, tais como congestionamentos, poluição e acidentes em massa.

O sistema de transporte brasileiro está baseado, principalmente, no transporte rodoviário, apresentando certo desenvolvimento no setor ferroviário, aéreo e aquaviário. Este último pouco aproveitado apesar da extensa capacidade fluvial do Brasil devido as suas características geográficas [12].

O processo de crescimento urbano ocorrido em meados do século 20 onde o Brasil permitiu que o país, predominantemente rural, transformasse-se em uma sociedade urbana intensa, com quase 80% de sua população morando em cidades. Essa mudança desencadeou uma intensa rotina de transporte de pessoas que passaram transferir-se de uma localidade a outra [13].

O transporte ferroviário teve início na Inglaterra por volta do século XIX devido principalmente a necessidade do transporte de matérias-primas até as fábricas e o retorno dos produtos manufaturados até os grandes centros urbanos. Após a II Guerra Mundial deu-se maior importância no desenvolvimento linhas metropolitanas e suburbanas e eletrificação das linhas em substituição aos motores a vapor e a diesel.

O transporte ferroviário, destaca-se pela sua capacidade de carga elevada, por sua eficiência energética e capacidade de transporte em grandes distâncias. No transporte de passageiros, este modal se destaca pela sua capacidade de transporte sem causar grandes congestionamentos, devido ao fato das locomotivas circularem de acordo com um sistema de tráfego controlado e com reduzida emissão de poluentes por utilizar o sistema elétrico como meio de alimentação. Considerando o problema de superlotação das capitais metropolitanas com grande concentração da população vivendo nas periferias e conseqüentemente o

transporte diário de pessoas entre o centro e os subúrbios, o transporte ferroviário se apresenta como boa solução para o problema, proporcionando a agilidade, devido a sua alta velocidade tráfego praticamente constante entre estações, sua pontualidade devido ao fato de não estar sujeita as variações no fluxo de veículos como ocorre no setor rodoviário e a capacidade de transporte muito superior àquelas permitidas pelo transporte rodoviário. O transporte ferroviário tem como característica básica o transporte de grandes massas. Diferentemente de outros modais ele não tem diferentes tipos de veículos com capacidades reduzidas de transporte, tal como os sistemas ferroviário e o aquaviários, compostos por veículos de grande e pequena capacidade, como por exemplo, ônibus e carros, navios e pequenas lanchas. O modal ferroviário é um meio de transporte capaz apenas de realizar o transporte em massa e com o aumento do número de passageiros o sistema ferroviário precisou evoluir a fim de suprir a demanda. Os sistemas de sinalização, largamente utilizados nas ferrovias brasileiras, têm como foco principal identificar se um trecho encontra-se ou não ocupado, não levando em consideração a velocidade do trem ao longo do trecho [18, 20] e conseqüentemente não permitindo um maior número de composições ativas simultaneamente.

As mudanças nos sistemas de sinalização precisaram ser implantadas levando em consideração os fatores de risco existentes, e para obter a máxima eficiência, mantendo a segurança o sistema, não pode basear-se somente no fator humano como foi desenhado inicialmente o sistema. Neste sistema, o maquinista era o único responsável por acelerar ou desacelerar o trem a partir das informações passadas pelos sinais existentes ao longo da via. [18]

Os sistemas de sinalização ferroviários estão cada vez mais automatizados e com níveis de controle cada vez mais apurados, por isso o desenvolvimento de novas formas de controle apresentam-se de forma bastante interessante para que o sistema mantenha-se seguro [18, 20]. Dentro deste contexto este trabalho tem por finalidade estudar uma forma de controle de velocidade totalmente independente dos elementos até então utilizados nas vias, baseada na detecção de movimento por meio de vídeos.

Utilizar câmeras ao longo das vias férreas permite o monitoramento do tráfego e a obtenção de diversas informações. Com o uso de câmeras ao longo da via é possível, por exemplo, detectar a velocidade do trem, classificar a frota em uso, acompanhar diferentes vias ao mesmo tempo, identificar e monitorar acidentes, detectar e avaliar pontos de falha, invasões, obstruções das vias, tudo isso sem que seja, necessariamente, alterado o sistema de sinalização em funcionamento.

Motivação

A intenção de utilizar a detecção de movimento para a estimação da velocidade de um trem surgiu da observação da necessidade de um sistema que pudesse aferir a velocidade independentemente daqueles relacionados ao circuito de via. Permitindo ao aos controladores do sistema ferroviário realizar a comparação entre as velocidades informadas pelo sistema em vigor, e aquelas obtidas com o sistema proposto.

As tecnologias disponíveis no mercado atualmente têm por objetivo aumentar a confiabilidade das funções de proteção, supervisão e operação do trem [18, 20], porém cada uma delas requer um alto investimento devido, principalmente, à tecnologia a ela agregada, à troca de diversos elementos ao longo da via férrea, dentro dos trens e nos sistemas de controle e monitoramento, os chamados CCOs (Centro de Controle Operacional).

No período de crescimento acelerado das grandes cidades, o investimento no sistema ferroviário brasileiro foi muito aquém daquele necessário para acompanhar a evolução tecnológica ocorrida neste setor. Os investimentos neste setor passaram a ser intensificados nos últimos anos com a privatização do sistema [7, 9] ainda assim muito modestos no que diz respeito à modernização dos sistemas, sendo a maior parte deste investimento foi realizado na manutenção do sistema [7]. Devido a isso os controles restringem-se as condições de identificação da proximidade dos trens e a ocupação da via, este tipo de análise não permite, por exemplo, determinar, de maneira precisa se um trem está ou não a uma velocidade superior àquela permitida e conseqüentemente não permite criar planos para o aumento do número de veículos na via.

O sistema por detecção de movimento informaria aos controladores de forma mais apurada a velocidade com que o trem atravessou determinado trecho independente da posição dele em relação ao circuito de via. O fato da velocidade aferida, por detecção de movimento, ser a velocidade instantânea permitiria assim ao controlador de maneira rápida solicitar a redução ou o aumento da velocidade por parte do maquinista para a correção de uma possível situação de risco.

A Figura 1 apresenta um diagrama de blocos da utilização do sistema proposto junto ao sistema de controle já existente e a utilização de suas informações no controle e monitoramento do tráfego.

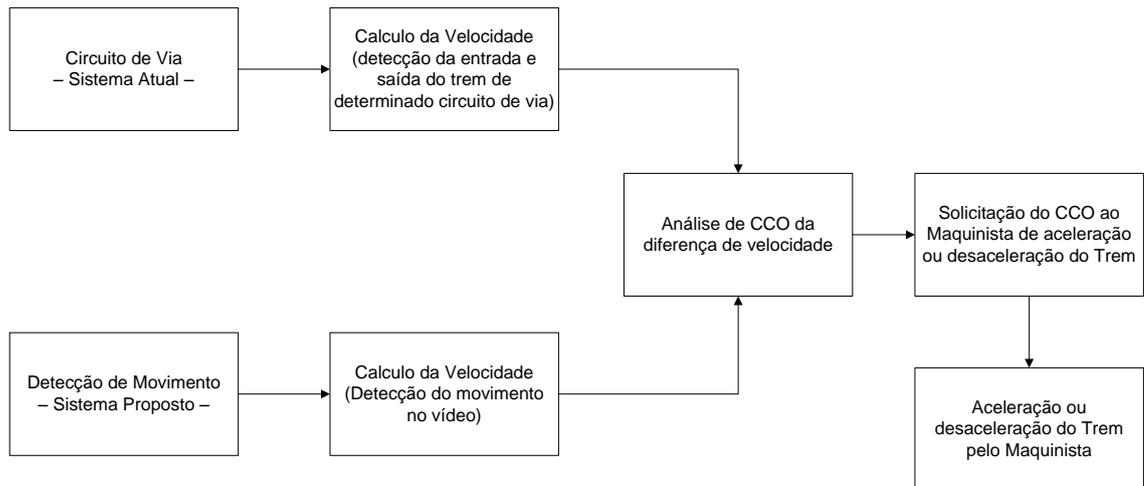


Figura 1. Diagrama de blocos da utilização do sistema

Objetivo

Esta tese objetiva estimar a velocidade dos trens utilizando o processamento de vídeos, para tanto foi desenvolvido um algoritmos com parâmetros de comparação que permitissem detecta o deslocamento do trem quadro a quadro.

As velocidades obtidas com tais algoritmos foram confrontadas com aquelas fornecidas pelo sistema de controle obtendo assim um parâmetro para validar os resultados obtidos pelo sistema em desenvolvimento nesta dissertação.

1 SISTEMA DE SINALIZAÇÃO FERROVIÁRIO

Os primeiros sistemas ferroviários baseavam-se na experiência e vigilância de seus condutores que precisavam manter-se atentos para detectar qualquer possível situação de perigo que solicitasse a parada do trem antes que este colidisse com o empecilho a sua frente, este sistema mostrou-se falho devido ao fato de nem sempre a distância necessária para o maquinista ver o obstáculo correspondia ao tempo preciso para que o trem parasse antes do obstáculo. A solução encontrada para que pelo menos um trem não se chocasse com o outro foi à utilização de intervalos de tempo entre as saídas dos trens, este intervalo compreendia o tempo necessário para que um trem saísse em velocidade máxima não encontrasse o trem anterior. Para que o condutor soubesse o exato momento de sua saída eram utilizados funcionários com bandeiras nas cores vermelha, amarela e verde, onde a sinalização com a primeira cor indicava que o condutor não poderia sair, a segunda que ele deveria preparar-se para a partida e a verde que este poderia dar início ao seu deslocamento [18]. Tal sistema, porém não permitia identificar possíveis problemas que pudessem ocorrer ao longo do percurso, tal como a quebra de um trem ou a obstrução da via por algum defeito dela, como por exemplo, o desalinhamento de um trilho, o sistema nestes casos permanecia totalmente dependente da percepção de seu maquinista. Uma solução adotada foi à divisão da via em trechos fixos onde em cada um deles só poderia movimentar-se um trem de cada vez. Para que fosse possível identificar se um trecho estava ou não ocupado eram instalados semáforos de sinalização, onde o semáforo vermelho indicava trecho ocupado e o verde trecho desocupado, conforme mostra a Figura 2.



Figura 2. Sinalização Fixa com sinal simples

Nestes casos, os trechos deveriam ser de um comprimento suficiente grande para que permitisse que o trem parasse antes de avançar sobre o trecho seguinte. Nos casos onde o trecho que contemplasse uma curva, o sinal usado do sinal simples poderia não ser suficiente,

visto que poderiam não permitir a visualização do semáforo do trecho seguinte, para estas situações fez-se necessária à utilização de sinais repetidores. Em casos, onde houvesse a necessidade de trens percorrendo o trecho com maior velocidade a solução foi à utilização de sinais triplos. A sinalização com semáforos triplos acrescentou ao sistema a cor amarela, utilizada para alertar ao condutor do trem que no trecho seguinte estaria ocupado, a Figura 3 exemplifica a utilização destes tipos de sinalização.

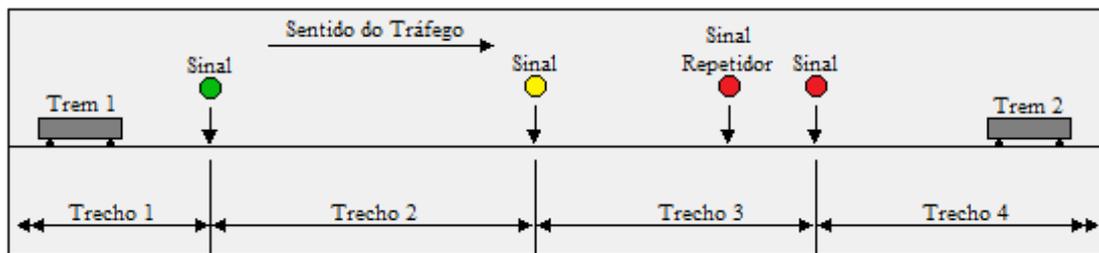


Figura 3. Sinalização Fixa com sinais triplos e repetidores

Toda a sistemática de sinalização anteriormente descrita foi, inicialmente, realizada através de sistemas mecânicos [18], operados por funcionários espalhados ao longo de toda a via. Os circuitos de via permitiram a detecção e sinalização automática dos trens por meio de relés eletromagnéticos fornecendo ao sistema maior confiabilidade.

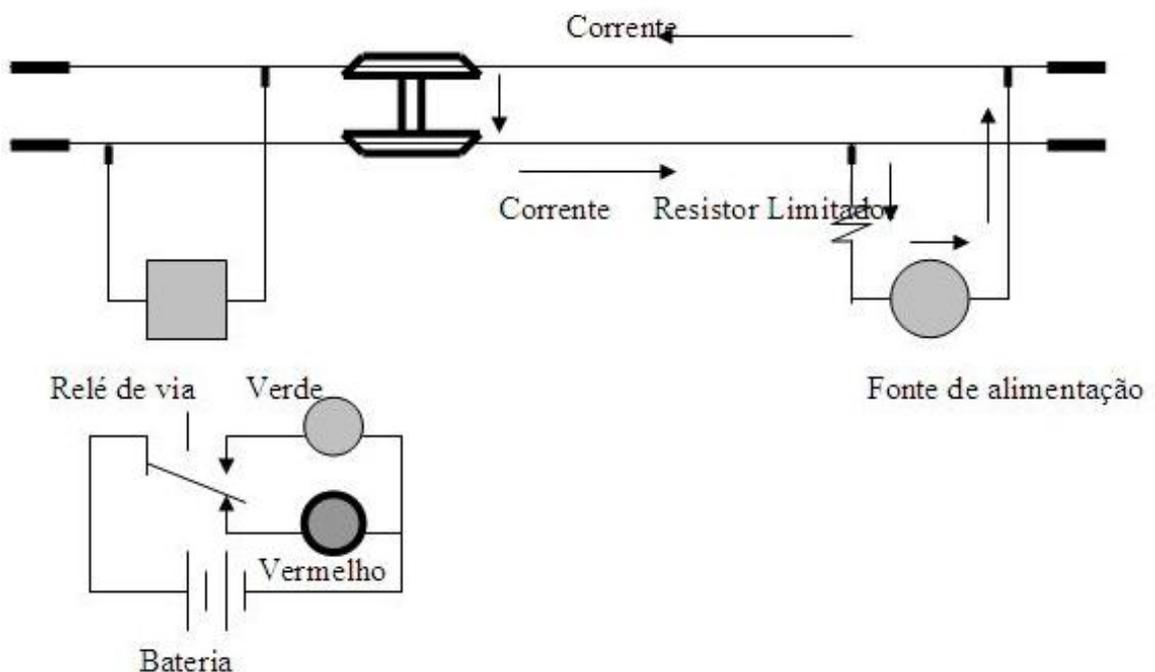


Figura 4. Circuito de Via

A automatização do sistema através da implementação do circuito de via ocorreu conforme ilustrada de forma simplificada a Figura 4[6].

Neste tipo de sistema os trilhos dos trechos adjacentes são isolados eletricamente em suas junções e as extremidades do trecho são conectadas uma a fonte de alimentação e a outra ao relé. Dessa forma o trem, ao entrar no circuito de via, acaba, devido a sua baixa resistência, por desviar a corrente elétrica fazendo com que o relé tenha sua alimentação interrompida, na ausência do trem, o relé é energizado pela fonte de alimentação. O fato de o relé estar ou não energizado implica na alteração das posições do contato, permitindo que este fique levantado ou abaixado [18]. Abaixado faz a lâmpada vermelha se acender, e levantado, faz a lâmpada verde se acender, esta configuração consegue prever que em casos de falha do sistema este ficará com a luz vermelha acesa, devido a este fato tal sistema é conhecido como vital ou de falha segura.

2 ANÁLISE E PROCESSAMENTO DOS VÍDEOS

Uma imagem digital é a codificação de uma cena usando números binários, onde cada ponto corresponde a uma determinada área da cena permitindo assim a reprodução, armazenamento, transferência e impressão. O menor elemento componente de uma imagem digital é denominado de “Picture elements” ou *pixel*. O processo de obtenção de uma imagem tem como principais parâmetros a frequência de amostragem espacial ou resolução espacial que determina o número de amostras obtidas por unidade de comprimento, vertical ou horizontal, do sensor e a quantização que permite determinar o número de bits representativos da intensidade de cinza ou cor de cada *pixel*, quanto maior o número de bits por *pixel*, maior será a taxa necessária para a transmissão até o aparelho de exibição de imagem e maior será a quantidade de memória necessária para armazenar a imagem, porém melhor será a qualidade da imagem.

O vídeo pode ser considerado a projeção de várias imagens em intervalos de tempo constante de forma que o olho humano tenha a ilusão de movimento. Nesse caso, cada imagem é chamada quadro ou quadros e a quantidade de imagens projetadas por segundo é dado o nome de quadros por segundo ou FPS (quadros per second). A qualidade do vídeo vai depender da quantidade de quadros capturados por segundo e da qualidade de cada quadro, que pode ser expressa pela quantidade de *pixels* utilizados (dimensão da tela) e da quantidade de informação em cada *pixel* (variação das cores).

O processamento de imagens é o processo automático de melhoramento, interpretação e análise quantitativa de uma imagem ou de um conjunto delas. Este processo consiste em transformar a imagem sucessivamente até que a informação objetivada seja dela satisfatoriamente extraída. Devido ao fato de ser uma atividade diretamente ligada ao sistema visual humano, esta torna-se complexa e dependente da interpretação da situação, devido a este fato não se pode definir um processo abrangente o suficiente para interpretar um grupo heterogêneo de situações, por isso a existência de diversos métodos de processamento foram desenvolvidos para atender a uma determinada necessidade.

Nem sempre, principalmente devido à complexidade do processamento, uma imagem é analisada por inteiro, em muitos casos trabalha-se com as regiões de interesse (ROI - Region of Interest). Tais áreas são normalmente definidas a partir de parâmetros previamente informados ao sistema. Através destes parâmetros podem-se identificar, na imagem, regiões onde o processamento deve ser intensificado e/ ou onde os resultados são de maior interesse.

O processo de segmentação da imagem na maior parte dos casos tem como foco inicial a retirada das informações pertencentes ao fundo da imagem. Esta prática é bastante adotada devido ao fato dos elementos de fundo em sua maioria conter um volume de informações muito alto, exigindo um tempo de processamento muito grande com resultados pouco relevantes ao processo. Existem muitas situações onde estes elementos inclusive deturpam os resultados obtidos, devido ao ruído que geram nas análises.

Em geral, não é possível capturar apenas as informações de interesse em uma imagem. Em detecção de movimento, como é o caso deste estudo, elementos secundários aos objetos analisados podem ter a mesma proporção daqueles de interesse o que pode atrapalhar o rastreamento do objeto, nestes casos é conveniente separar, ou quando possível retirar tais informações para que extraído e utilizando uma menor quantidade de dados para análise, resultado seja um conjunto de áreas de interesse com o menor volume possível de informações desnecessária para a análise objetivada.

O processo de reconhecimento de padrões é fundamental para identificar semelhanças existentes entre as imagens, este processo visa classificar os elementos de uma imagem segundo um conjunto de parâmetros pré-definidos. A partir deste reconhecimento é possível iniciar as análises quantitativas aplicando ferramentas específicas à situação a ser avaliada.

O sistema desenvolvido é baseado no processamento de imagens de vídeo capturadas através de uma câmera. Os vídeos obtidos são todos coloridos, porém as imagens obtidas no processo são convertidas para imagens em tons de cinza, devido à complexidade das imagens coloridas. Tal conversão permite que o processamento seja mais rápido visto que em imagens coloridas seria necessário realizar as comparações nas três dimensões enquanto na escala de cinza isto é feito em apenas uma. Apesar de haver ganho no processamento com esta conversão perde-se precisão do processo, visto sua identificação em um quadro posterior seria mais assertiva do que aquela utilizando a informação de apenas uma dimensão. As imagens capturadas são importadas para o MATLAB sendo convertidas para uma matriz de pontos com valores entre 0 e 255 que representam a intensidade luminosa de cada ponto.

2.1 Movimento

O movimento real de objetos no espaço ocorre em três dimensões. Em alguns casos pode-se avaliar o movimento que ocorre em três dimensões utilizando apenas duas delas realizando uma projeção da terceira, isto pode acontecer quando o deslocamento ocorrer de

forma mais significativa em apenas dois eixos do percurso. Neste estudo será realizada a aproximação para o deslocamento em apenas duas dimensões. Esta aproximação pode ser feita, primeiro, devido ao fato do deslocamento do objeto em uma das dimensões ser desprezível em relação às outras. Segundo devido ao posicionamento do dispositivo de captura que permitiu a sobreposição dos elementos a fim de que todos eles estivessem no mesmo eixo de visão.

2.2 Detecção de Movimento

Um sistema de detecção de movimento tem por definição a capacidade de identificar a existência de movimentos em uma determinada sequência de imagens. Isto é realizado em função de um conjunto de parâmetros previamente informados determinados de acordo com a utilização e as características do sistema a ser monitorado. As técnicas utilizadas na detecção do movimento podem ser aplicadas, por exemplo, a sistemas de vigilância, análise de movimento humano, rastreamento de pedestres e veículos, análise sanguínea e até mesmo abertura e fechamento de portas [2,14].

A detecção de movimento tem por objetivo final conceder às máquinas e aplicações à capacidade de não somente captar e armazenar imagens [2,14]. Pretende-se com a detecção de movimento obter do sistema monitorado informações quantitativas e objetivas, analisando-o no intuito de obter informações de maneira mais ágil quando comparada ao método utilizado pela visão humana. Além da agilidade obtida com a análise computacional do movimento retira-se do sistema a parte subjetiva vinculada às avaliações humanas, trazendo maior confiabilidade aos resultados obtidos.

Um algoritmo considerado eficaz deve ser prático, com bom tempo de resposta, robusto a presença de ruídos e preciso. Dependendo de sua aplicação algumas destas características devem ser destacadas em relação a outras. Isto porque dependendo da aplicação do algoritmo a velocidade da resposta pode, por exemplo, ser mais crítica que a sua precisão.

Existem basicamente dois grupos de sistemas básicos para detecção do movimento, um deles avalia um conjunto de características, tais como bordas e/ou cantos, informados previamente, para que o sistema rastreie o objeto a ser analisado [2,14]. No outro sistema é verificada a variação luminosa da imagem identificando onde ocorreu a maior variação é possível responder o que foi movimentado. Dentro destas linhas de raciocínio pode-se ainda

analisar imagens consecutivas ou comparar uma imagem com uma imagem intitulada imagem modelo. Ambos possuem aspectos positivos e negativos, sendo sua aplicação dependente diretamente do sistema a ser analisado.

O algoritmo que utiliza uma imagem modelo, ou imagem de fundo, precisa levar em consideração alguns parâmetros para a definição de seus elementos imóveis, visto que dependendo do tempo de análise uma pessoa parada por alguns instantes pode ser considerada erroneamente como parte do fundo. Uma solução intermediária seria a comparação entre quadros adjacentes, porém nos casos onde os elementos movem-se lentamente, a variação entre quadros nesse caso seria tão pequena que poderia por acabar caracterizando-a como um ruído.

Já o sistema que realiza sua comparação através da variação luminosa sofre bastante interferência em locais não confinados, ou seja, em locais ao ar livre, visto que um dia de chuva com muitos raios, ou até mesmo a sombra gerada pelo novo posicionamento do objeto poderia inviabilizar o sistema como um todo.

Outro fator a ser levado em consideração é o processamento utilizado nestes algoritmos para se analisar o movimento. Por exemplo, uma imagem com resolução de 640 x 480 *pixels* com 256 níveis de cinza, 8 bits de profundidade por *pixel*, com taxa de amostragem de 30 quadros por segundo solicitaria um processamento de mais de 73 Mbits/seg. Assim é importante não só aplicar o algoritmo, mas conseguir determinar a informação essencial a ser processada para que o sistema possa atender no tempo necessário e a um custo compatível a realidade do projeto.

Uma solução para esta situação é a realização de buscas de movimento em blocos de *pixels*, ao invés de em *pixels* individuais. Essa estratégia permite reduzir o tempo de processamento sem diminuir a confiabilidade do sistema [2, 14]. Algoritmos que verificam todas as posições possíveis de um quadro são conhecidos como algoritmos de busca completa. Uma maneira de otimização de detecção de movimento é a limitação do campo de busca. Neste caso ao invés de realizar a busca do bloco ao longo de todo o quadro o algoritmo realiza a busca em uma determinada janela, definida de acordo com o padrão de movimento determinado pelo objeto e pela capacidade de processamento permitida pelo projeto, tais algoritmos são conhecidos como algoritmos de busca rápida [2, 14].

2.3 Redução de Ruído

Os vídeos obtidos possuem variações na intensidade dos *pixels* devido a diferenças ocorridas no momento de sua captação. Essas alterações aleatórias são classificadas como ruídos [13], elas alteram o valor dos *pixels* comprometendo significativamente a qualidade das fotos. O fato de tais variações ocorrerem por flutuações na cor e/ou na luminância, no caso do sistema proposto, pode levar o processo de comparação a uma identificação incorreta. Isto porque, como todo o processo de detecção de movimento, aqui utilizado, realiza a diferença entre a intensidade de *pixels* entre um conjunto de *pixels* de um bloco no conjunto de *pixels* do quadro seguinte, um conjunto destes em um quadro pode ter um valor diferente da mesma região no quadro seguinte o que dificultaria o processo de detecção utilizado. Uma técnica utilizada para amenizar o problema é a utilização de filtros [5].

O processo de filtragem de imagens consiste no realce de uma determinada faixa de frequências, normalmente classificadas em altas, médias e baixas. A filtragem acontece através da utilização de matrizes denominadas máscaras. Este ocorre através da passagem destas máscaras ao longo de toda a imagem analisando o conjunto de *pixels* englobados pela matriz escolhida, processo conhecido como convolução.

A filtragem, através da convolução, é um processo difundido e de simples aplicação que opera no domínio espacial. A definição do tamanho das máscaras e seus elementos são fatores determinantes neste processo para que a operação pretendida seja bem sucedida [5].

Os filtros passa - baixa, são os mais comumente usados para o realce de imagens, esses têm como característica a suavização das altas frequências, evitando assim transições abruptas. Seus efeitos nas imagens são a minimização do ruído e a suavização da mesma [5].

Os filtros gaussianos podem se comportar como filtros passa-baixa, visto que para cada ponto analisado é calculada a intensidade média de alguns pontos vizinhos, este processo diminui assim as intensidades das componentes em altas frequências. A Equação 1 descreve a equação para implementação de um filtro Gaussiano.

$$h_{(n_1, n_2)} = \frac{e^{-\frac{(n_1^2 + n_2^2)}{2\sigma^2}}}{\sum_{n_1} \sum_{n_2} e^{-\frac{(n_1^2 + n_2^2)}{2\sigma^2}}} \quad (1)$$

Na qual n_1 e n_2 são as dimensões da máscara a ser aplicada e σ a variância.

Outro processo bastante difundido para melhora na nitidez das imagens consiste em filtrar a imagem original utilizando um filtro passa baixa. Esta imagem suavizada é subtraída da imagem original gerando assim uma máscara, conhecida por máscara de nitidez. Esta máscara é adicionada a imagem inicial com um determinado peso. Acrescentar pesos positivos é o processo conhecido por filtragem de alto reforço (High-boost), enquanto que pesos negativos atenuam os resultados obtidos para a máscara de nitidez [5]. As equações que expressa à máscara de nitidez pode ser expressa pela Equação 2.

$$g_{máscara}(x, y) = f(x, y) - f_f(x, y) \quad (2)$$

Na qual $f(x,y)$ e $f_f(x,y)$ e correspondem à imagem original e a filtrada respectivamente e $g_{máscara}(x,y)$ refere-se à máscara de nitidez.

A ponderação realizada na máscara pode ser descrita através da equação 3.

$$g(x, y) = f(x, y) + k * g_{máscara}(x, y) \quad (3)$$

Na qual $f(x,y)$ corresponde à imagem original, $g_{máscara}(x,y)$ refere-se à máscara de nitidez e k a ponderação.

Os filtros não lineares são filtros espaciais muito uteis na eliminação de ruídos impulsivos, dentro desta categoria de filtros um filtro bastante utilizado é o filtro de medianas. Este filtro consiste em substituir o valor de um *pixel* centrado em uma determinada vizinhança pela mediana das intensidades dos seus vizinhos. Comparando este tipo de filtragem com as filtrações lineares pode-se perceber que nesta o efeito de borramento da imagem é menor [5 7].

Os três filtros descritos foram utilizados no processamento da imagem antes do inicio dos algoritmos de detecção de movimento. A comparação entre os resultados obtidos pode ser observada no capítulo de discussão de resultados.

2.4 Métrica de Similaridade

A métrica de similaridade é responsável por determinar quais blocos são mais ou menos similares dentro do universo de busca estipulado. Devido ao fato destes critérios permitirem identificar o quão diferente uma região é relativamente à outra, eles podem também ser chamados de critérios de distorção. Dentre as diversas métricas desenvolvidas pode-se destacar o erro médio absoluto, o erro quadrático médio e a soma da diferença absoluta [19].

A soma da diferença absoluta (*SAD*, *Sum of Absolute Differences*) dentre os métrica de similaridade apresentados é o mais simples e o mais utilizado, principalmente devido ao fato do número de operações a serem executadas quando usado ser menor do que para outras métricas, reduzindo assim o nível de processamento [19].

Esta métrica calcula a distorção entre o bloco atual e o bloco de referência, utilizando a diferença absoluta entre eles, esta diferença é realizada *pixel a pixel* e sua função é dada pela Equação 4:

$$SAD = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |C_{ij} - R_{ij}| \quad (4)$$

Na qual, C_{ij} corresponde ao bloco do quadro em análise, R_{ij} ao bloco de referência e M e N as dimensões do bloco.

O erro absoluto médio (*MAE*, *Mean Absolute Error*) calcula o valor médio do módulo da diferença entre os *pixels* do bloco atual com os *pixels* do bloco de referência, resultando no erro absoluto das regiões pesquisadas conforme mostra a equação 5:

$$MAE = \frac{1}{M * N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |C_{ij} - R_{ij}| \quad (5)$$

Com as variáveis iguais as descritas na Equação 4.

O erro quadrático médio (MSE, *Mean Absolute Error*), dentre os aqui citados, é a métrica que requer a maior complexidade de processamento devido à potenciação executada no resultado da diferença entre os blocos, a equação que representa o MSE é a Equação 6:

$$MSE = \frac{1}{M * N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (C_{ij} - R_{ij})^2 \quad (6)$$

Novamente com as variáveis iguais as descritas pela equação 4.

Vale neste caso destacar que nenhuma métrica aqui descrita obtém resultados negativos, visto que em todos os cálculos da diferença obtida ou é elevado ao quadrado ou é calculado seu valor absoluto, evitando assim que erros negativos anulem erros positivos.

A métrica utilizada neste estudo foi o SAD, devido a menor complexidade de processamento.

3 SISTEMA PROPOSTO

Este trabalho tem por objetivo avaliar diferentes métodos de detecção de movimento por meio de vídeos para aferir a velocidade de um trem circulante em uma via férrea. Partindo do princípio de que é possível estimar, através da comparação de quadros de um vídeo, o deslocamento de um determinado objeto presente nele, e que todos os vídeos, utilizados neste estudo, possuem uma taxa de amostragem constante, o processo de determinação da velocidade passaria a ser uma divisão simples entre o deslocamento observado e o tempo necessário para que tal deslocamento ocorra, a Figura 6 apresenta uma visão sistêmica do processo.



Figura 5. Visão do Sistema Proposto

3.1 Características do Sistema

O sistema de aquisição de imagens utilizado para obter os vídeos com a passagem dos trens é composto de uma câmera digital de vídeo, um computador e o software utilizado para a estimativa da velocidade através da detecção do movimento.

A câmera digital utilizadas no processo foi uma Canon modelo PowerShot A570IS com CCD colorido de 1/2.5 polegadas e aproximadamente 7.1Mpixels, lentes de distância

focal 5.8(W) – 23.2(T) mm e zoom óptico e zoom digital de aproximadamente 4x. Os vídeos foram capturados a uma taxa de 30fps e as dimensões do quadro são de 640x480 *pixels*. O computador utilizado foi um notebook da marca Dell com processador Intel Core i5 M460 de 2.53GHz e 4GB de memória RAM. Os vídeos foram todos importados para o programa chamado MATLAB, versão R2008a. Este foi utilizado como ferramenta no processo de rastreamento e alteração das imagens.

3.2 O Banco de Dados

Em cada local visitado foram realizadas medições para identificação da distância da via ao dispositivo de captura, e realizados pelo menos dois vídeos do trem movimentando-se em cada configuração. Além da disponibilidade de acesso, a escolha dos locais foi determinada por outros dois aspectos. No primeiro levou-se em consideração os aspectos dinâmicos do local, o objetivo era obter o fundo o mais estático possível. O segundo aspecto levado em consideração, principalmente, regiões onde a velocidade do trem no trecho não sofresse alterações bruscas de velocidade, tais como entradas e saídas de estações, locais próximos a passagens de nível, entre outros.

Os dias, horários e locais onde as imagens foram capturadas foram anotados, conforme Tabela 1. Através deles foi possível realizar o levantamento da velocidade junto ao centro de controle. Nem todos os vídeos coletados puderam ter suas velocidades obtidas, aqueles onde foi possível tal levantamento foram utilizados para as análises iniciais onde os resultados encontrados puderam ser comparados àqueles fornecidos pelos sistemas vigentes.

Nº Arquivo	Horário de Passagem	Circuito de Via	Nº da Via	Distância T - M (m)	Altura da Camera (m)	Sentido	Tamanho do CV (m)
MOV01357	02:07	189T	2	21,2	1,2	E-D	238
MOV01358	02:08	189T	3	17,5	1,2	E-D	251
MOV01359	02:13	189T	4	13,3	1,2	D-E	249
MOV01361	02:16	189T	2	21,2	1,2	D-E	238
MOV01362	02:19	189T	3	17,5	1,2	E-D	251
MOV01363	02:22	189T	4	13,3	1,2	D-E	249
MOV01364	02:24	189T	1	24,9	1,2	E-D	198
MOV01367	02:39	189T	2	21,2	1,2	E-D	238
MOV01368	02:46	189T	4	13,3	1,2	D-E	249
MOV01369	02:48	189T	4	13,3	1,2	D-E	249
MOV01370	02:50	189T	3	17,5	1,2	E-D	251
MOV01371	02:51	189T	1	24,9	1,2	D-E	198

Tabela 1. Levantamento dos trens

3.3 Parâmetros de Captura dos Trens

A realização das tomadas foi feita com a câmera digital disposta a uma distância definida do chão, através do uso de um tripé. Tomou-se o cuidado de manter o dispositivo paralelo ao solo de forma que este não ficasse na diagonal para que assim pudesse ser evitadas alterações devido aos ângulos formados, variou-se também a distância entre o trem e o dispositivo de captura, conforme mostra a Figura 6.



Figura 6. Distâncias entre o dispositivo e os objetos

A distância entre o dispositivo e o veículo ferroviário é importante para a conversão do deslocamento em *pixel* para metros. Esta conversão leva em consideração algumas características da câmera fotográfica utilizada, no caso deste estudo uma câmera digital.

O funcionamento de uma câmera digital é bastante semelhante em certos aspectos ao funcionamento de uma câmera tradicional, que utiliza filme fotográfico, em vias gerais o processo de captura é bem semelhante, as diferenças mais significativas iniciam no processo de armazenagem da imagem. Enquanto em máquinas comuns a armazenagem ocorre em filmes com propriedades químicas que alteram de acordo com a exposição deste a luz, nas máquinas

digitais o CCD (*Charge-Coupled Device*) é o responsável por transformar as intensidades de luz que incidem sobre ele em valores digitais.

O processo de formação da imagem na câmera ocorre conforme apresentado na Figura 7. Este ocorre através de um conjunto de lentes posicionadas tal forma a apresentarem o comportamento semelhante ao de uma lente convergente, a imagem é formada no filme, no caso das câmeras comuns ou no CCD nos casos das câmeras digitais.

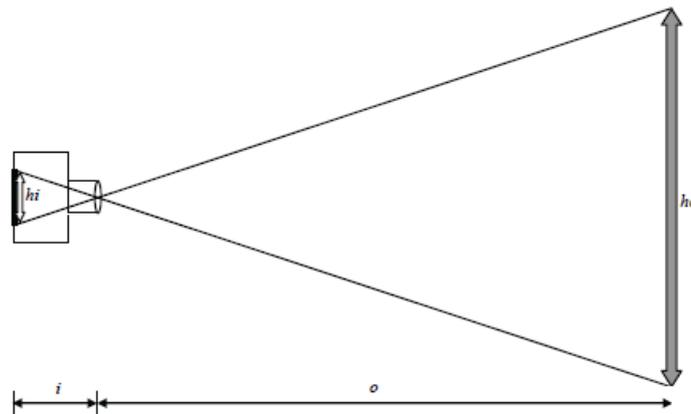


Figura 7. Captura da Imagem

A lei de Gauss permite calcular a distância i entre a média de altura h_i e o centro óptico da lente, conforme equação 7, onde f é a distância focal, o a distância da lente ao objeto:

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i} \quad (7)$$

Ou

$$i = \left(\frac{1}{f} - \frac{1}{o} \right)^{-1} \quad (8)$$

Pela figura 7 é possível identificar dois triângulos semelhantes formados pelos catetos h_o , o e h_i , i e por isso possível escrever a Equação 6:

$$\frac{h_o}{o} = \frac{h_i}{i} \quad (9)$$

Neste estudo, o foco é o cálculo da altura, h_o do objeto, logo pode-se reescrever a equação como:

$$h_o = \frac{h_i}{i} o \quad (10)$$

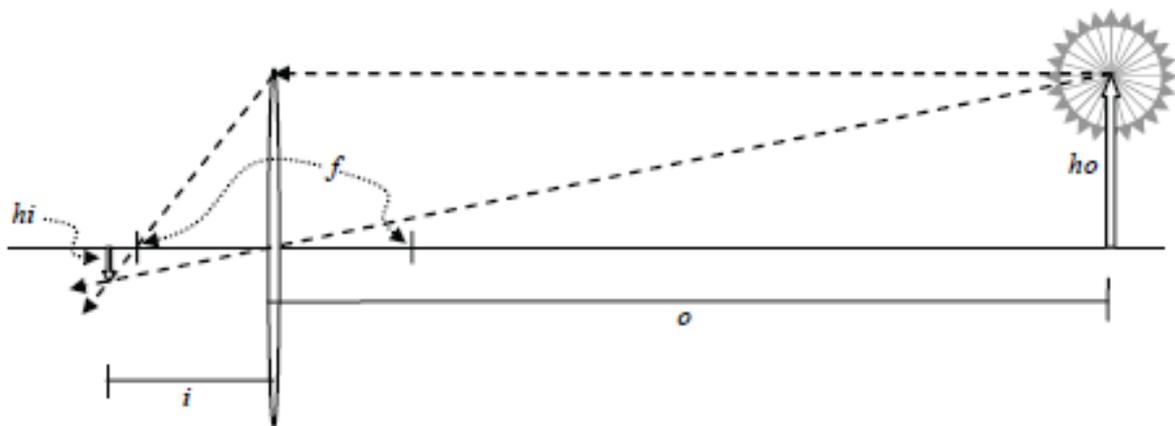


Figura 8. Formação de uma imagem

A equação descrita permite assim o cálculo do comprimento real do objeto capturado, conhecendo este valor e sabendo as dimensões em *pixels* do quadro, é possível através de uma regra simples converter o número de *pixels* em metros. O deslocamento em metros pode ser escrito então como:

$$D_m = \frac{D_p * h_o}{h_p} \quad (11)$$

Na qual D_m é o deslocamento em metros, D_p é o deslocamento em *pixels*, h_p é a altura total do quadro em *pixels* e h_o é a altura do objeto em metros.

3.4 Imagem de Referência

O processo inicial para estimar o movimento em vídeos é o processo de classificação dos elementos que compõem o fundo e os elementos que compõem os objetos em movimento. O sistema proposto tem como primeiro passo a identificação do quadro inicial onde é possível detectar o trem, ou seja, para otimizar o processo de cálculo da velocidade desconsidera-se o conjunto de quadros onde o trem ainda não está presente no vídeo. O processo de detecção de movimento é feito comparando a imagem inicial, ou imagem de referência, com as imagens seguintes, conforme Figura 9.

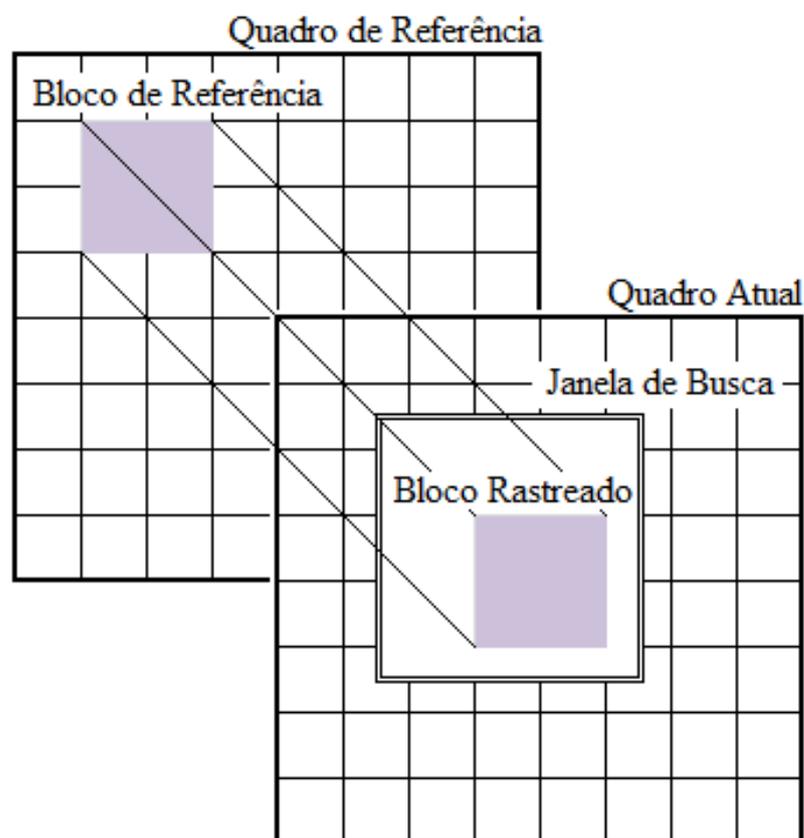


Figura 9. Detecção do Movimento

Nota-se que no processo de comparação dos quadros as imagens onde não se observa a presença do trem são muito semelhantes, imagens por isso classificadas como imagens pertencentes ao segundo plano ou imagens do fundo. Devido a essa grande similaridade os valores encontrados para a diferença entre os conjuntos de *pixels* destas imagens é bastante reduzida, e aumenta conforme o trem vai aparecendo nos quadros seguintes. As pequenas

diferenças obtidas no início dos vídeos podem ser atribuídas ao ruído de captação dos vídeos, tais como movimentações de objetos leves devido à brisa do local de captura e vibrações no solo devido à aproximação do trem. O mesmo processo pode ser observado no momento em que o trem começa a sair, neste caso observa-se que a diferença retorna a valores muito próximos daqueles obtidos com os quadros iniciais.

Esta comparação entre quadros é bastante eficaz, neste caso, devido a dois principais motivos. As regiões de captura de o vídeo possuir pouquíssimos elementos em movimento, devido às características de uma via férrea, que, por motivos de segurança, solicitam que a região de circulação de trens deva sempre buscar ser uma área com o menor número de elementos dinâmicos permitidos para evitar possíveis acidentes. Esta condição, porém nem sempre pode ser obtida devido à configuração urbana atual. As regiões vizinhas às linhas ferroviárias tem apresentado grande concentração de pessoas. Para evitar que esta situação polua o vídeo com elementos dinâmicos o local para a captura deles deve ser o mais distante possível destes locais de grande concentração urbana, para que assim a predominância de elementos em movimento seja a de trens.

O outro motivo favorável à detecção de movimento através da comparação entre quadros são as dimensões do trem. Não é necessário que o objeto de captura esteja muito próximo do veículo ferroviário para que este ocupe grandes regiões do quadro, sendo por isso facilmente detectado. Por esta análise fica evidente que a distância entre o dispositivo de captura e o trem influencia diretamente na detecção do fundo da imagem, quanto maior a região ocupada pelo trem, mais fácil à detecção de seu aparecimento no vídeo.

O processo de determinação do conjunto de quadros a serem avaliados no processo de detecção de movimento utilizou a comparação entre quadros consecutivos. Isto é, o algoritmo utilizados comparou, dois a dois, todos os quadros pertencentes ao vídeo, esta comparação foi realizada entre o quadro anterior e o seguinte. A soma da diferença absoluta entre todas as intensidades dos *pixels* contidos entre estes dois quadros foi armazenada e comparada. No momento em que esta diferença aumentou significativamente foi classificada com o início da movimentação do trem ao longo do vídeo. O quadro correspondente a maior diferença foi então utilizado como o quadro inicial para o processo de detecção. A mesma linha de raciocínio foi utilizada para a determinação do último quadro a ser utilizado na detecção do movimento. O passo seguinte foi determinar o limiar de transição entre o quadro com movimento e o quadro sem movimento, o parâmetro utilizado foi realizar a média entre os valores obtidos anteriormente e posteriormente aquele ponto. Caso a média diferísse aquele caso fosse a primeira variação seria considerado o primeiro quadro contendo o trem e caso

fosse o último a classificação seria de último quadro contendo o trem. A Figura 10 apresenta o gráfico obtido com o processo acima de descrito, nela é possível identificar que o quadro onde o trem iniciou seu processo de passagem foi logo após o quadro de número 150, concluindo logo após o quadro de número 500. Com este exemplo é possível perceber que dos mais de 600 quadros que o vídeo possui o trem está presente em aproximadamente 350 quadros deles. Logo a utilização deste método de detecção, para o exemplo apresentado na Figura 5, permitiu a redução de 42% do número de quadros a serem analisados nas etapas seguintes.

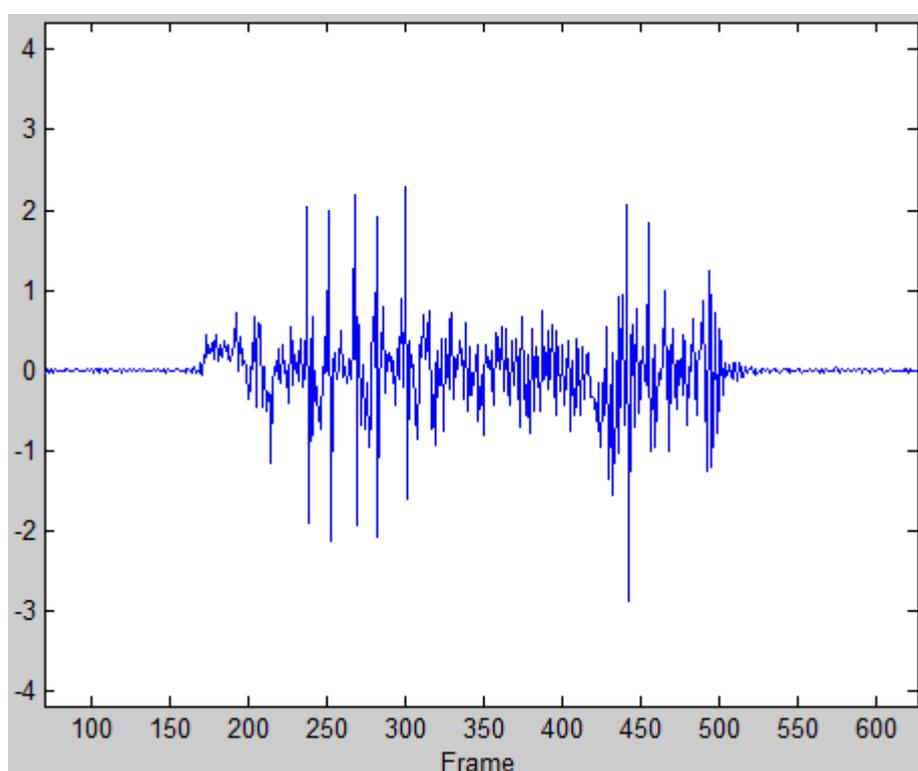


Figura 10. Comparação entre quadros

3.5 Tamanho da Janela

Em geral, ao realizar uma filmagem é muito difícil conseguir um enquadramento que contenha apenas as informações a serem analisadas. Regiões de informações pertencentes ao fundo aumentam o tempo de processamento e podem gerar distorções no resultado obtido.

Na análise de deslocamento dos trens ao longo da via férrea é possível constatar que o deslocamento do trem ocorre basicamente na horizontal. O deslocamento na vertical ocorre apenas em trechos em declive ou alicive, nos casos de defeitos ao longo da via, e devido às

vibrações que possam ocorrer devido às características de montagem e acoplamento dos vagões que compõem o trem. Isto porque o movimento abrupto em um vagão propaga-se aos vagões a ele conectados. Considerando tais características de deslocamento, regiões superiores e inferiores dos quadros poderiam ser descartadas sem que houvesse perda de informação relevante para o sistema. O processo de obtenção da janela ideal para o sistema foi definida observando a região de passagem dos trens ao longo das vias nos vídeos coletados. Desta observação identificou-se que uma janela com uma altura de 128 *pixels* centrada no meio do trem conseguiria analisar apenas o conjunto de *pixels* pertencente à região de passagem de trem no vídeo.

Devido grande comprimento do trem em relação ao campo de visão do dispositivo de captura para que o dispositivo conseguisse enquadrar todo o trem na largura do seu quadro seria necessário que este dispositivo estivesse muito distante da via o que não foi possível devido às restrições do local de obtenção das amostras. O distanciamento do dispositivo permitiria o enquadramento, porém aumentaria muito a quantidade informações desnecessária na análise do movimento. Todos os vídeos obtidos estavam a uma distância tal que o comprimento total do trem era sempre maior do que a largura do quadro, por isso optou-se por uma janela de largura igual à largura do quadro para que assim todo o movimento do trem pudesse ser utilizado no sistema.

3.6 Tamanho e posicionamento dos blocos

Outro parâmetro a ser dimensionado no sistema são os blocos de referência e comparação. O formato do bloco escolhido neste sistema foi o quadrado, tal formato foi escolhido por ser de fácil definição de coordenadas.

O tamanho do bloco de comparação depende muito das características dos *pixels* presentes no quadro. Se a imagem analisada possuir uma quantidade elevada de *pixels* com valores de intensidade muito próximos uns dos outros, o resultado das comparações entre blocos será muito próximo podendo acarretar muitos erros na determinação do bloco mais similar. Neste caso é interessante utilizar blocos com dimensões maiores para que um maior conjunto de *pixels* possa ser comparado, apesar do aumento das dimensões do *pixel* reduzir o número de comparações realizadas ao longo da janela de busca.

Nos casos onde as imagens capturadas apresentam *pixels* com grande variação em sua intensidade luminosa às comparações baseadas na diferença entre estes valores apresentarão

valores relativamente mais distintos o que permitiria a utilização de blocos menores, porém a redução da dimensão do bloco aumenta o número de comparações aumentando o nível de processamento.

Ao analisar alguns quadros foi possível perceber que em diversas regiões ao longo dele eram observadas repetições de padrões, tais como portas, janelas e partes da lataria, o que permitiu identificar que blocos pequenos teriam maior chance de errarem o processo. Sendo assim optou-se por determinar a melhor dimensão através de tentativas com diferentes tamanhos para definição daquele que resultaria em um sistema mais robusto. Para tanto foram testados blocos de dimensões 8x8, 16x16, 32x32, 48x48, 64x64 e 128x128 *pixels*, os resultados obtidos com cada um dos blocos será comentado no capítulo de discussão de resultados.

A localização do bloco de referência na janela também deve ser levada em consideração, visto que um bloco colocado na parte esquerda do bloco realizaria o maior número de comparações possíveis, pois iria deslocar-se ao longo de todo o comprimento do vídeo, pois o processo de comparação iniciaria no momento em que o trem iniciasse seu deslocamento no vídeo. Esse tipo de análise só estaria correta nos casos onde o deslocamento do vídeo ocorresse da esquerda para a direita, caso o deslocamento ocorresse no sentido inverso, além do início do processo ocorrer apenas quando a frente do trem já estivesse se deslocado ao longo de todo o comprimento do vídeo, está não conseguiria estimar o movimento, visto que o bloco mais semelhante já estaria fora do quadro seguinte e por isso sem possibilidade de detecção pelo sistema.

A solução encontrada para evitar tais problemas foi colocar o bloco de referência bem no meio da janela para que este pudesse obter o deslocamento em qualquer direção em que o trem se locomovesse. Este tipo de abordagem, porém não permite que o rastreamento seja realizado ao longo de todo o quadro apenas em parte dele.

3.7 O Rastreamento

O processo de detecção de movimento inicialmente utilizado, procurava ao longo do quadro áreas específicas do trem como, por exemplo, janelas, portas, frisos. Estas regiões depois de identificadas eram rastreadas ao longo do vídeo, a Figura 11 apresenta um quadro com a detecção da janela com o friso vermelho de um trem. O ponto demarcado representa a localização do centróide do retângulo contendo janela e o friso vermelho.



Figura 11. Padrão de uma janela identificada

Esta linha de análise apresentou bons resultados, porém se mostrou pouco eficiente devido à basicamente dois motivos: Primeiro, pode-se destacar a heterogeneidade da frota de trens em uso no Brasil, segundo devido a prática de utilização da lataria área externa do trem serem atualmente utilizadas a colocação de grandes faixas de propaganda de diferentes empresas que são atualizadas periodicamente. Devido aos motivos apresentados ficou claro que a busca por elementos comuns a todos os trens seria uma tarefa impossível. A Figura 12 apresenta alguns dos modelos de trens e seus diferentes padrões.



Figura 12. Padrões dos Trens

Optou-se então por um sistema mais robusto que não rastreasse um padrão pré-determinado, o processo então passou a buscar um determinado bloco não necessariamente vinculado a um padrão pertencente ao trem. O processo de rastreamento comparou dois a dois os blocos do quadro de referência com os blocos do quadro atual. Foram utilizados três tipos distintos de comparação, no primeiro caso utilizou-se um bloco de referência estático, no segundo um bloco de referência dinâmico e no terceiro, um bloco cujo valor de intensidade

dos *pixels* era definido pela média das intensidades dos *pixels* dos blocos anteriormente definidos. Cada um dos métodos apresentados será detalhado nas subsecções a seguir.

3.7.1 Bloco Estático

Neste método, um bloco de dimensões $M \times M$ *pixels*, posicionado conforme seção 3.6, é utilizado como bloco de referência do quadro N para a comparação com os k blocos do frame seguinte, onde N varia entre 1 e o número de quadros total do vídeo. Após esta comparação um novo processo de comparação se inicia entre o quadro $N+1$ e o quadro $N+2$. O bloco utilizado como bloco de referência no quadro $N+1$ tem dimensões e coordenadas exatamente iguais àsquelas utilizadas para definir o bloco do quadro N . O bloco nesta abordagem é então chamado de bloco estático devido ao fato de sua posição não ser alterada com a mudança dos frames. A Figura 13 apresenta um exemplo ilustrativo de como acontece este tipo de abordagem. Nela, o bloco destacado em laranja é utilizado como bloco de referência para a comparação entre o frame N e $N+1$. Em lilás está representado o bloco de referência para a comparação entre os frames $N+1$ e $N+2$. Os blocos em vermelho representam os blocos identificados nos quadros seguintes com similares.

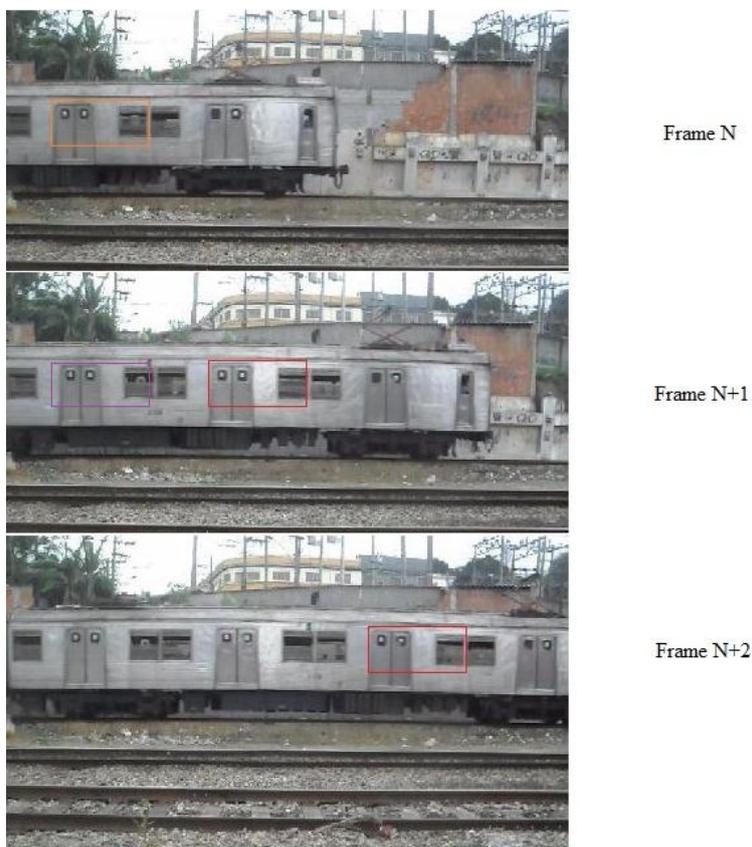


Figura 13. Sistema com bloco de referência estático

3.7.2 Bloco Dinâmico

O segundo método comparativo utilizado tem como princípio a utilização de blocos de referência com posições que variam à medida que são alteradas as duplas de quadros de comparação. Devido a essa alteração constante no posicionamento do bloco de referência este bloco foi intitulado bloco dinâmico.

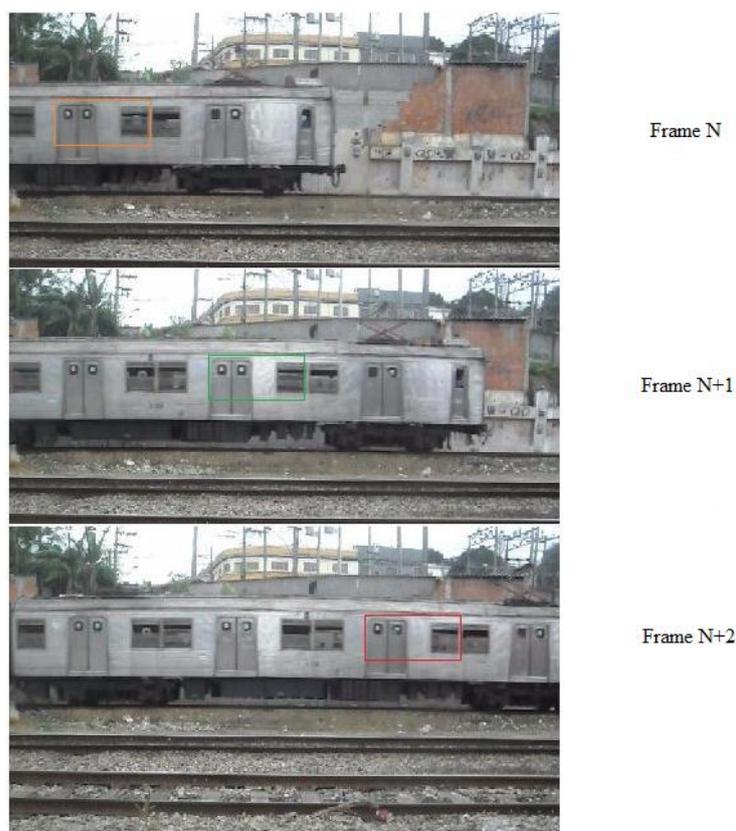


Figura 14. Sistema com bloco de referência dinâmico

A alteração no posicionamento do bloco, não foi realizada de forma aleatória, o cálculo do novo posicionamento foi realizado com base na identificação do bloco mais similar obtido na comparação da dupla de quadros anterior. Isto é, um quadro N com um bloco de dimensões $M \times M$ posicionado, inicialmente, de acordo com a secção 3.6 é utilizado como bloco referencial para a comparação com todas as possíveis combinações de blocos de mesma dimensão $M \times M$ obtidos no frame N+1. A métrica de similaridade descrita na secção 2.4 é utilizada com todos estes blocos. Dentre estas comparações um bloco é eleito como o mais semelhante àquele descrito no quadro N. Este bloco, pertencente ao quadro N+1 e localizado

em uma posição de coordenadas (x_1, y_1) , é utilizado como bloco de referência para as comparações a serem realizadas entre os quadros $N+1$ e $N+2$. A Figura 14 ilustra o procedimento adotado. Nela, um bloco referencial de cor laranja é utilizado para comparação com os blocos do quadro $N+2$, o bloco identificado como o mais semelhante é representado pela cor verde. Este bloco de cor verde por sua vez é utilizado para as comparações a serem realizadas com o quadro do bloco $N+2$, o bloco obtido nesta comparação está representado pela cor vermelha.

Por fim o terceiro método de rastreamento utilizado foi o de cálculo de médias de intensidades dos *pixels* do bloco de referência. O processo inicial neste algoritmo é exatamente igual àquele realizado nos métodos anteriores, ou seja, um bloco escolhido em uma determinada posição é procurado no quadro seguinte. Ao obter o bloco do quadro seguinte é realizada a média das intensidades entre estes dois blocos. O novo bloco de referência formado é composto pela média das intensidades dos *pixels* dos dois quadros anteriores. A cada novo bloco obtido, são considerados os valores dos *pixels* dos blocos obtidos anteriormente, com isso espera-se que o sistema seja menos sensível a possíveis ruídos que possam ocorrer ao longo do vídeo. Espera-se com isso que este algoritmo seja mais robusto que os anteriores.

3.8 Comparação entre Blocos

O critério de comparação utilizado para determinar quais blocos são mais ou menos similares dentro do universo de busca estipulado, foi o de soma da diferença absoluta. Neste, é realizada a diferença entre cada *pixel* do bloco de referência com o *pixel* correspondente no bloco analisado. O módulo da diferença é calculado e o somatório de todas as diferenças obtidas do conjunto de *pixels* dos blocos é armazenada. Este processo é repetido ao longo de toda a janela de análise comparando com todas as combinações de blocos possíveis. O bloco cuja diferença tenha o menor valor absoluto é considerado o bloco mais semelhante ao bloco de referência.

Tomando como base as matrizes apresentadas na Figura 15 para identificar o bloco no quadro de busca que mais se assemelha ao bloco de referência, obtêm-se os seguintes valores de SAD em cada uma das interações $SAD1 = 26$, $SAD2 = 36$ e $SAD3 = 25$. Neste exemplo então o bloco com maior nível de similaridade é o terceiro e àquele com o menor é o de índice 2.

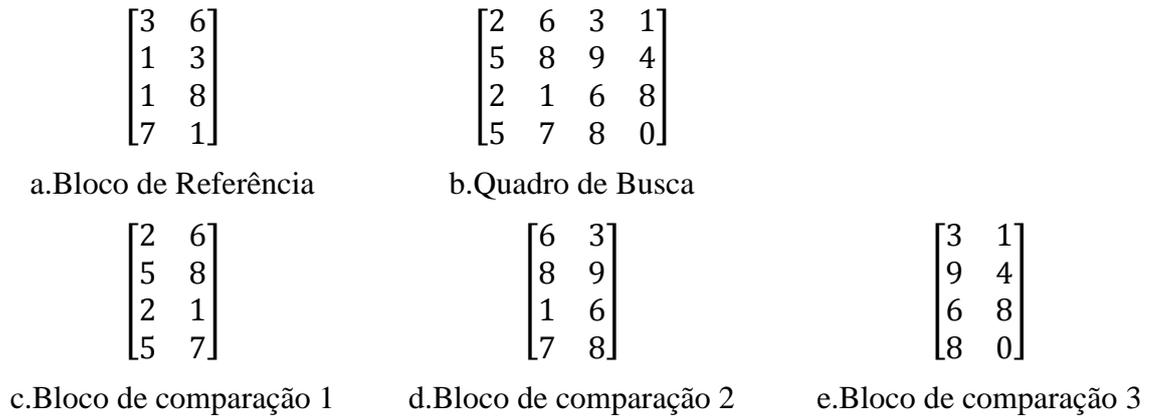


Figura 15. Comparação entre blocos

3.9 Determinação da Velocidade

O cálculo da velocidade de passagem do trem pela via férrea, objeto deste estudo, será realizado baseado no deslocamento de um objeto em relação a uma câmera estática. Este deslocamento pode ser considerado bidimensional se projetada no plano da imagem. [9]. Desta forma estamos obtendo uma medida de velocidade média, porém levando em consideração o curto espaço percorrido pelo bloco e o tempo de captura menor que 1 segundo, os valores estimados podem ser aproximados para a velocidade instantânea.

Vale ressaltar que nos casos onde foi possível realizar o levantamento das velocidades dos trens através do centro de controle às capturas foram realizadas no menor circuito de via disponível para que a velocidade fornecida pelo sistema sofresse a menor variação possível. Isto porque, para o sistema em utilização, a velocidade é calculada apenas considerando a diferença entre o horário que o trem ocupou o circuito e o horário que este foi liberado, ou seja, a velocidade média no trecho, não contemplando as possíveis variações de velocidade que o maquinista possa ter realizado.

O deslocamento obtido através da estimativa do movimento é obtido em *pixel* permitindo assim o cálculo da velocidade em *pixels* por quadro, neste caso é necessário converter esta velocidade em metros por segundo. Esta conversão é possível considerando que o quadro inteiro possui uma largura fixa definida pelo próprio dispositivo e que através da Equação 7 é possível determinar a largura do quadro em metros, a conversão do deslocamento em *pixels* para metros é, então, determinada pela equação 12:

$$D_m = \frac{h_o D_p}{L_p} \quad (12)$$

Na qual D_m corresponde ao deslocamento em metros, D_p o deslocamento em *pixels*, h_o a largura total do quadro em metros e L_p a largura total do quadro em *pixels*.

A conversão de quadros em minutos pode ser obtida através da velocidade de captura do dispositivo de captura fornecida em quadros por segundo. Como todas as considerações de deslocamento são inferidas entre dois quadros adjacentes, tempo necessário para que ocorra o deslocamento é igual ao inverso da velocidade de captura. O cálculo da velocidade em metros por segundo pode ser definida pela equação 13.

$$V_m = \frac{D_m}{V_c^{-1}} \quad (13)$$

Na qual D_m corresponde ao deslocamento em metros e V_c a velocidade de captura do vídeo.

3.9.1 Validação dos Testes

A obtenção das velocidades dos trens não foi possível em todos os vídeos capturados, isto se deve em parte pela disponibilização das informações e outra pela limitação dos sistemas em vigor. O processo inicial de confronto dos resultados obtidos com a sistemática nesta tese descrita, e aquela utilizada pelos sistemas em uso atualmente foi realizada utilizando este grupo com velocidade conhecida. Esta fase inicial ficou definida como fase de validação dos testes, os vídeos seguintes foram utilizados para confronto de métodos utilizados e dos melhoramentos aplicados a técnica seja na imagem gerada seja na otimização dos resultados.

3.9.2 Estimativa de Erro

A métrica utilizada na validação dos resultados foi a do erro quadrático, esta compara os resultados de velocidade média informados pelo sistema de controle atualmente em uso e a velocidade obtida pelo sistema proposto descrito pela Equação 14:

$$e = (V_{esp} - V_{calc})^2 \quad (14)$$

Na qual e corresponde ao erro quadrático, V_{esp} a velocidade esperada e V_{calc} a velocidade obtida pelos algoritmos.

A velocidade média informada pelo sistema de controle vigente na empresa em questão, conforme anteriormente descrito, refere-se a uma velocidade média calculada em função do tempo de permanência do trem do circuito de via. Devido a isso ao longo deste trecho o trem pode ter variado sua aceleração, por isso o valor da velocidade média referente ao circuito de via seguinte torna-se importante para identificar o comportamento do trem no trecho.

Os vídeos obtidos para a validação dos testes foram realizados próximos ao final do circuito de via. Esta condição permitiu comparar as velocidades médias obtidas ao longo dos trechos adjacentes da via férrea. Esta comparação faz-se necessária visto que trens com variações de velocidade significativas ao longo do trecho acabariam por gerar uma velocidade média muito diferente daquelas obtidas com o sistema proposto, onde a velocidade média obtida compreende um trecho muito menor do que aquele utilizado pelo sistema de controle ferroviário.

As chances de, em trechos reduzidos, o comportamento de a aceleração variar muito são bem menores quando comparadas aquelas ocorridas em trechos maiores, por isso as estimativas de erro sempre levarão em consideração os valores de velocidades informados nos trechos adjacentes para avaliação do comportamento do trem no trecho em questão.

Outra estimativa de erro considerada é aquela obtida entre os valores de velocidade média obtida entre trechos adjacentes. Este erro apresenta a imprecisão existente na metodologia de obtenção da velocidade utilizada pelo sistema em funcionamento atualmente, a ser calculada através da Equação 15.

$$e_{var} = (V_{esp} - V_{trecho})^2 \quad (15)$$

Na qual e_{var} corresponde ao erro quadrático, V_{esp} a velocidade esperada e V_{trecho} a velocidade média do trecho adjacente àquele onde V_{trecho} foi obtida.

4 RESULTADOS

As velocidades obtidas neste capítulo representam o módulo da velocidade estimada. O sentido do movimento é dado pelo sinal, onde o sinal positivo representa um movimento no sentido positivo do eixo x, ou seja, da esquerda para a direita, e o sinal negativo representa um movimento no sentido negativo, da direita para a esquerda. O sentido de deslocamento do trem é de fácil detecção e não é um fator muita importância e por isso nos ateremos principalmente ao módulo da velocidade.

Inicialmente, serão apresentados os resultados obtidos a partir do tratamento da imagem para a redução de ruídos comparando o tempo de processamento e os resultados obtidos com aqueles obtidos sem o tratamento da imagem. Em seguida, serão apresentados os resultados dos diversos métodos de estimação da velocidade dos trens a partir de vídeos vistos no capítulo 3. A apresentação dos métodos aplicados está de acordo com a cronologia do desenvolvimento dos métodos e por isso os primeiros métodos são mais simples e os últimos mais complexos.

4.1 Comparação dos resultados obtidos para os filtros

O método de filtragem Gaussiana apresentou os mesmos resultados de erros independentes da variação de seus parâmetros, porém o tempo de processamento entre eles variou ligeiramente. Isto acontece devido às características da imagem, visto que esta apresenta grandes regiões similares, na região do trem as grandes variações ocorrem nas regiões de fronteira entre elementos, ou seja, na mudança parte metálica dele para uma região de porta ou janela.

A tabela 2 compara o tempo de processamento e o erro na estimação da velocidade entre a imagem com redução de ruído e a imagem sem a utilização de um filtro.

Os dados mostram que o erro obtido no processo é ligeiramente menor quando a imagem é tratada por um filtro Gaussiano de máscara de dimensões 3x3 e variância igual a 0,5. Este resultado era esperado, pois, como visto na seção 3.4, tais filtros têm como principal característica a suavização da imagem o que permitiria ao sistema obter blocos mais parecidos, porém o fato do resultado não ser muito superior é razoável devido às características de uniformidade da imagem anteriormente comentadas.

Tipo de Filtro	Parâmetros	Erro	Tempo de Processamento (s)
Sem filtro		1,88%	539,00
Gaussiano	Máscara [3 3] Variância 0.5	1,87%	503,24
Gaussiano	Máscara [5 5] Variância 0.5	1,87%	524,54
Gaussiano	Máscara [3 3] Variância 0.8	1,87%	509,75
Gaussiano	Máscara [3 3] Variância 3	1,87%	502,34

Tabela 2. Resultados obtidos com o filtro Gaussiano

O processo combinado de máscara de nitidez e filtragem *high-boost*, assim como o filtro Gaussiano, apresentou melhora nos resultados obtidos quando comparado ao processo realizado sem o tratamento da imagem, porém este tipo de filtragem, dentre aqueles analisados, foi o processo que demandou maior tempo de processamento.

Tipo de Filtro	Parâmetros	Erro	Tempo de Processamento (s)
Sem filtro		1,88%	539,00
Máscara de Nitidez	K=1	1,92%	758,48
Máscara de Nitidez	K=3	1,89%	660,41
Máscara de Nitidez	K=5	2,48%	617,94

Tabela 3. Resultados obtidos com a máscara de nitidez e Filtragem *High-Boost*

Tal resultado era esperado, visto que o número etapas é maior, como apresentado na seção 3.4. A etapa inicial, de obtenção da máscara de nitidez utilizou o próprio filtro Gaussiano, o que suporta a consideração de aumento de processamento. A tabela 3 apresenta os resultados obtidos para este tipo de tratamento.

O método de filtragem utilizando o filtro de medianas resultou nos melhores resultados dentre os métodos de redução de ruídos aqui avaliados, sendo este obtido utilizando uma máscara de dimensões 3x3 conforme apresentado na Tabela 4. Este tipo de filtro conseguiu reduzir os *pixels* errados gerados aleatoriamente na imagem que poderiam não aparecer em um quadro e aparecer no quadro seguinte, dificultando assim ao processo de detecção de deslocamento dos blocos obter o bloco correto.

A melhora obtida com este filtro, porém resultou na redução do erro encontrado no processo de apenas 0,02%, com o aumento um aumento de 6% no tempo de processamento do algoritmo.

Tipo de Filtro	Parâmetros	Erro	Tempo de Processamento (s)
Sem filtro		1,88%	539,00
Filtro de Medianas	Máscara [3 3]	1,88%	515,14
Filtro de Medianas	Máscara [5 5]	1,86%	569,94

Tabela 4. Resultados obtidos com o Filtro de Medianas

Devido a pouca melhora obtida com as técnicas de filtragem aqui expostas e o aumento significativo do tempo de processamento causado por elas, optou-se por retirar tal etapa do processo utilizando para todos os métodos avaliados a imagem originalmente obtida pelo dispositivo de captura.

4.2 Bloco único estático

Este método foi apresentado na subseção 3.7.1. E a seguir discutiremos os seus resultados com diferentes tamanhos de bloco de referência.

4.2.1 Bloco de 8x8 pixels

Com blocos de referência de 8x8 obtemos uma velocidade média muito acima da esperada, mais de 100% de aumento na velocidade. Analisando os dados obtidos pelo algoritmo percebemos que este apresenta um conjunto de aproximadamente apenas 30% do total de quadros, que correspondem à passagem do trem, nos quais a velocidade estimada está próxima à velocidade esperada do trem. Os outros quadros, que juntos representam 70% do total de quadros com trem, resultam em velocidades diferentes da esperada, 20% do total de quadros resultaram em velocidades muito superiores à esperada e 50% retornaram velocidades próximas de zero. Como a velocidade esperada é relativamente próxima de zero se comparada às altas velocidades encontradas a velocidade média resultante pairou acima do dobro da velocidade esperada. O erro de estimação, apresentado na seção 3.9.2, ficou em $e=1126,38\%$.

Tamanho do Bloco	Tempo de Processamento (segundos)
8x8	584,81
16x16	564,05
32x32	537,46
48x48	522,55
64x64	503,63
128x128	22,88

Tabela 5. Tempo de processamento obtidos para o bloco único estático com diferentes tamanhos

Foi constatado que esse efeito se deve à possibilidade de existência de um bloco muito semelhante ao bloco de referência que possa causar a confusão no cálculo da métrica de semelhança entre blocos, base do sistema de estimação da velocidade do trem. A grande

semelhança entre blocos se deve primeiramente à uniformidade da estrutura do trem, os vagões têm sua lataria praticamente idêntica em todos os pontos, com poucas exceções como: portas, janelas e pintura não homogênea ou defeituosa. Nas regiões onde a lataria é uniforme o cálculo do SAD, que é calculado sobre a imagem em escala de cinza, resulta em valores muito próximos. O segundo fator que impele o algoritmo ao erro é o tamanho do bloco, como será visto nas subseções 4.1.2 a 4.1.5, que conforme se aumenta bloco diminui-se a incidência de erro. Isso ocorre porque blocos maiores contêm informações de uma superfície maior, o que permite que particularidades desta superfície tornem-se mais aparentes, fazendo com que estes blocos passem a apresentar características inconfundíveis, portanto tornando-se mais fáceis de serem detectados.

Outro fator a ser considerado é o tempo de estimação da velocidade, para blocos de 8×8 pixels de tamanho o processo durou 585 segundos. Foi o processo mais demorado dentre os processos testados, conforme a Tabela 5.

O número de vezes que um SAD é calculado para um determinado tamanho de bloco pode ser obtido com a Equação 16:

$$N_{SAD} = (L_j - L_b + 1)(H_j - H_b + 1), \quad (16)$$

Na qual L_j e H_j são respectivamente largura e altura da janela e L_b e H_b são as dimensões do bloco.

Neste caso, pode-se notar que quanto maior o bloco menor o número de SADs a serem calculados. Por outro lado o SAD tem um número de operações que é dado pela Equação 17:

$$O_{SAD} = 2L_b H_b - 1 \quad (17)$$

Com as variáveis iguais as descritas na Equação 16.

Neste caso quanto maior o bloco, maior o número de operações, o que é intuitivo. Assim, a combinação das equações 16 e 17 resulta em um número total de operações de:

$$N_{SAD} O_{SAD} = (L_j - L_b + 1)(H_j - H_b + 1)(2L_b H_b - 1) \quad (18)$$

Com as variáveis iguais as descritas na Equação 16.

Analisando o comportamento desta equação chegaremos à conclusão que quanto maior o bloco maior o número de operações e, portanto mais demorado deveria ser o processo. Porém, durante a execução do algoritmo, no momento da varredura da janela, a cada vez que o SAD é calculado entre o bloco de referência e o bloco em análise, há um carregamento do bloco em análise, sendo esta operação relativamente lenta. Por sua vez o número de carregamentos é igual ao número de SADs calculados, NSAD, o que faz com que blocos de tamanhos menores, que têm um valor de NSAD maior, levem um pouco mais de tempo para serem calculados em relação a blocos maiores.

Portanto, o tempo efetivo de processamento e, principalmente, a quantidade de erros obtidos torna o bloco de 8x8, que tem como principal atrativo o baixo custo computacional para o cálculo do seu SAD, inviável. A tabela 6 mostra os resultados obtidos para o bloco de tamanho 8x8 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
8x8	1126,38%	584,81	50	20

Tabela 6. Resultados médios obtidos para o Bloco único Estático de 8x8

4.2.2 Bloco de 16x16 pixels

Dobrando cada uma das dimensões do bloco de 8x8 *pixels*, ou seja, aumentando em quatro vezes o número de elementos avaliados obtemos um resultado melhor do que o resultado com um bloco menor, porém continuamos a ter incidência de erros na detecção da posição de certos blocos, o que conforme explicado anteriormente na seção 4.2.1., se deve à semelhança entre os blocos existentes no quadro avaliado. A taxa de incidência de erro com blocos desse tamanho é de 55%, sendo que 52% dos resultados obtidos correspondem a baixas velocidades e os 3% restantes a altas velocidades, o que produziu um erro de

estimativa de $e=41,89\%$. O tempo necessário para a estimação da velocidade com blocos desse tamanho é de 564 segundos, pouco menos do que os blocos de 8×8 *pixels* de tamanho. Novamente temos um bloco de configuração inadequada uma vez que a taxa de erro impossibilita a estimação fidedigna da velocidade. A tabela 7 mostra os resultados obtidos para o bloco de tamanho 16×16 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
8x8	1126,38%	584,81	50	20
16x16	41,89%	564,05	52	3

Tabela 7. Resultados médios obtidos para o bloco único estático de 16×16

4.2.3 Bloco de 32×32 *pixels*

Como era de se esperar um novo aumento do tamanho do bloco produziu uma redução na taxa de incidência de erro que passou a ser de 32% do total de quadros onde há a presença do trem. Desta vez não houve erros que produzissem velocidades acima de 50% da velocidade esperada. O erro de estimação da velocidade ficou para este caso em $e=70,20\%$. Mesmo com a redução na quantidade de erros eles ainda representam 30% da totalidade dos valores de velocidade retornados, o que inviabiliza este tamanho de bloco para a estimação da velocidade do trem. Se comparado ao erro de estimação do bloco de 16×16 , visto na subseção 4.2.2, o erro do bloco de 32×32 é superior, isso ocorre porque nesta configuração não existem valores discrepantes em velocidades muito superiores à velocidade estimada para balancear a média de velocidade. Um fato relevante é que o bloco de 32×32 *pixels* corresponde a uma parcela bem grande da superfície da janela avaliada e mesmo assim há uma grande incidência de erros, o que aponta para a grande homogeneidade da lataria do trem. O tempo de processamento de blocos de 32×32 foi de 537 segundos e obedece ao comportamento esperado. A tabela 8 mostra os resultados obtidos para o bloco de tamanho 32×32 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
8x8	1126,38%	584,81	50	20
16x16	41,89%	564,05	52	3
32x32	70,20%	537,46	31	0

Tabela 8. Resultados médios obtidos para o Bloco único Estático de 32x32

4.2.4 Bloco de 48x48 pixels

Com blocos de 48x48 *pixels* obteve-se uma taxa de incidência de erros discrepantes na estimativa de velocidade de 2%, estes erros correspondendo a velocidades inferiores à metade da velocidade esperada. Esse resultado aponta para a existência de características únicas nos blocos avaliados. O erro de estimação da velocidade ficou em $e=6,75\%$. Este resultado era esperado uma vez que a superfície abrangida pelo bloco é bem grande. Deste modo podemos considerar que a partir desse tamanho de bloco temos um retorno fiel da velocidade real por parte do algoritmo. Além disso, o seu tempo de processamento é de 523 segundos o que é menor do que o tempo de processamento dos blocos menores o que torna sua utilização ainda mais atrativa. A tabela 9 mostra os resultados obtidos para o bloco de tamanho 48x48 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
8x8	1126,38%	584,81	50	20
16x16	41,89%	564,05	52	3
32x32	70,20%	537,46	31	0
48x48	6,75%	522,55	2	0

Tabela 9. Resultados médios obtidos para o bloco único estático de tamanho 48x48

4.2.5 Bloco de 64x64 pixels

Blocos de 64x64 ocupam a metade dos *pixels* de altura da janela e 10% da largura desta, ou seja, o bloco ocupa 5% da superfície total da janela. Com este tamanho de bloco as taxas de erros discrepantes caem para 0%, ou seja, não há erros de detecção. O erro de estimação da velocidade por sua vez fica em $e=3,55\%$. Nota-se que um bloco deste tamanho produz resultados muito precisos, não está sujeito a erros e tem tempo de processamento de 504 segundos. Devido aos ótimos resultados obtidos com esse bloco ele será o bloco de referência para os testes posteriores. A tabela 6 mostra os resultados obtidos para o bloco de tamanho 64x64 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
8x8	1126,38%	584,81	50	20
16x16	41,89%	564,05	52	3
32x32	70,20%	537,46	31	0
48x48	6,75%	522,55	2	0
64x64	3,55%	503,63	0	0

Tabela 10. Resultados médios obtidos para o bloco único estático de tamanho 64x64

4.2.6 Bloco de 128x128 pixels

Estes blocos ocupam 20% da área da janela, sendo que sua altura é a mesma da janela. Seguindo as expectativas de melhora do resultado conforme o aumento do tamanho do bloco era de se esperar um resultado superior ao encontrado com blocos de 64x64 *pixels*. Porém, o erro de estimação de velocidade desse método fica em $e=11352,54\%$, erro este que inviabiliza

o seu uso. Esse erro ocorre porque um bloco com 128 *pixels* de altura passa a englobar também uma parte do fundo seja este céu, o muro que restringe o acesso às vias, o solo ou qualquer outro objeto que possa fazer parte do fundo. E como esses objetos são estáticos eles não acompanham a região do trem que se deseja detectar, e por isso interferem no cálculo do SAD. A taxa de erro discrepante neste caso é de 100%, o que indica que a presença do fundo nos cálculos faz com que o algoritmo não detecte o bloco procurado. Devido as dimensões do bloco o tempo de cálculo da velocidade por esse método é bem pequeno quando comparado aos anteriores, sendo igual 23 segundos. A tabela 11 mostra os resultados obtidos para o bloco de tamanho 128x128 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
8x8	1126,38%	584,81	50	20
16x16	41,89%	564,05	52	3
32x32	70,20%	537,46	31	0
48x48	6,75%	522,55	2	0
64x64	3,55%	503,63	0	0
128x128	11352,54%	22,88	100	0

Tabela 11. Resultados médios obtidos para o bloco único estático de tamanho 128x128

4.3 Dois Blocos Estáticos

Este método foi apresentado nas subseções 3.6. e 3.7.1 e é muito semelhante ao método do bloco único fixo a não ser pela presença de um segundo bloco para estimação da velocidade. Neste caso os blocos têm seu ponto médio horizontal e vertical alinhado com o ponto médio vertical da janela, e cada um é disposto afastado horizontalmente de uma distância d do meio da janela. Assim temos os dois blocos em posições simétricas em relação ao eixo vertical. Neste método calculam-se duas velocidades, cada uma relativa a um bloco,

sendo o resultado final igual a média das duas. Por estimar a velocidade de dois blocos este método é duplamente suscetível aos erros de estimativa discrepantes.

Outra peculiaridade deste método é a possibilidade do uso de métricas entre os blocos para a determinação da velocidade. Uma vez que distâncias, como a distância entre os centróides dos blocos ou distâncias entre determinados pontos dos elementos podem ser facilmente calculadas, pode-se fazer estimativas conjuntas de velocidades. A seguir serão mostrados os resultados obtidos com blocos de 48x48 *pixels* e 64x64 *pixels*. A tabela 12 apresenta os tempo de processamento necessários para cada um dos blocos testados.

Tamanho do Bloco	Tempo de Processamento (seg)
8x8	616,72
16x16	595,74
32x32	567,66
48x48	990,77
64x64	950,11

Tabela 12. Tempos de processamento obtidos para dois blocos estáticos com diferentes tamanhos

4.3.1 Bloco de 48x48 *pixels*

Com esse tamanho de bloco conseguiu-se um erro de estimativa de $e=2,08\%$, com taxa de erro de valores discrepantes de 1,5%. Um desses erros trouxe um alto valor de velocidade e por isso a média foi elevada, aproximando-a do valor de referência. A ocorrência dos erros de valores discrepantes causa uma perda significativa da confiabilidade do resultado, pois pode ocorrer desses erros compensarem as diferenças reais entre a velocidade de referência e a velocidade medida aproximando o erro de estimação de zero. O tempo de processamento para esta configuração foi de 991 segundos. A tabela 13 mostra os resultados obtidos para o bloco de tamanho 48x48 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
48x48	2,08%	990,77	2	1

Tabela 13. Resultados médios obtidos para dois blocos estáticos de tamanho 48x48

4.3.2 Bloco de 64x64 pixels

O erro de estimação neste caso foi de $e=3,87\%$ sem incorrer em erros discrepantes. Nota-se que neste caso a velocidade de um bloco é, em média, ligeiramente inferior à outra velocidade obtida pelo outro bloco. O comportamento entre a velocidade de um bloco e a velocidade do outro bloco não é igual quadro a quadro, indicando a detecção de dois movimentos distintos para um mesmo objeto, o trem. Este comportamento não apresenta diferenças muito acentuadas entre as velocidades, conforme mostra a figura 16.

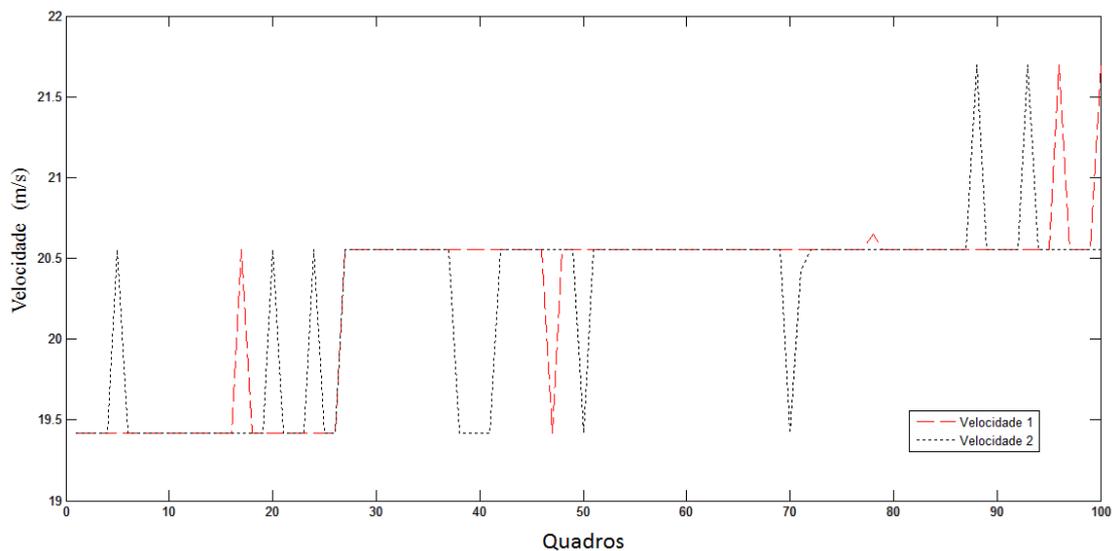


Figura 16. Velocidade dos blocos estáticos ao longo do vídeo

Isso indica que a escolha do bloco em diferentes posições afeta a estimação da velocidade, mesmo que infimamente. O que nos sugere a existência de um bloco ótimo para a estimação da velocidade de um objeto deslocando-se em um vídeo. Os resultados foram obtidos depois de 950 segundos de processamento. A Tabela 14 apresenta os valores obtidos para o bloco de 64x64 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
48x48	2,08%	990,77	2	1
64x64	3,87%	950,11	0	0

Tabela 14. Resultados médios obtidos para dois blocos estáticos de tamanho 64x64

4.4 Bloco Único Dinâmico

Visto nas seções 3.6. e 3.7.2., este método consiste no rastreamento de um único bloco, inicialmente posicionado no centro da janela, deslocando-se ao da janela de busca. Uma vez que este bloco de referência, presente em um quadro N, sair da janela de busca em um quadro N+1; outro bloco de referência é escolhido no centro da janela no quadro N, evitando quadros adjacentes onde não haja cálculo de velocidade ou que esta resulte em um valor absurdo por erro de detecção. A escolha deste novo referencial ocorre quando uma relação entre o número de quadros por segundo, o tamanho do bloco avaliado e a posição atual do bloco, atinge certo patamar, indicando o tempo de permanência máxima de um determinado bloco na janela.

Aqui podemos destacar três tipos de erro de estimação principais. O erro de estimação leve, que não é propagado, que consiste em uma detecção de um bloco, B_e , que tem uma grande região em comum com o bloco que deveria ser encontrado, B_t , e resulta em um novo bloco de referência, $B_{ref}=B_e$, alguns *pixels* deslocado da posição onde ele deveria estar

causando um leve aumento ou redução pontual da velocidade estimada do trem. Por assumir uma posição próxima ao bloco que deveria ser detectado o bloco $B_{ref}=B_e$ faz parte do trem e por isso não causa problemas em estimações futuras.

O erro de estimação médio ocorre quando o bloco encontrado B_e está distante do bloco B_t , mas é composto por uma região semelhante à região que deveria ser detectada e por isso causa um grande aumento ou redução pontual na velocidade estimada e não causa nenhum problema nas estimações futuras. Esse erro corresponde ao erro de estimação discrepante.

O erro de estimação grave pode levar o método a resultados errôneos. Este tipo de erro ocorre quando o algoritmo não consegue detectar um bloco B_e , nas proximidades de B_t ou semelhante a este, o que gera um grande erro e desloca o SAD do novo referencial para um valor muito diferente do qual este deveria assumir, desencadeando uma sequência de resultados errados. Estes resultados só voltarão a ser fiéis ao deslocamento do trem quando o tempo de permanência do bloco na janela se esgotar e o bloco de referência for reposicionado na origem. A Tabela 15 mostra os tempos de processamento obtidos com este método.

Tamanho do Bloco	Tempo de Processamento (seg)
8x8	732,26
16x16	706,45
32x32	672,21
48x48	829,75
64x64	809,44

Tabela 15. Tempos de processamento obtidos para o bloco único móvel com diferentes tamanhos

4.4.1 Bloco de 48x48 pixels

Nesse caso notou-se uma taxa de erros discrepantes de 4% e um erro de estimação de $e=2,54\%$. Notou-se que o erro de estimativa da velocidade foi reduzido, pois existem erros que, mesmo sendo esparsos, devido ao seu alto valor elevam à média. Com blocos desse tamanho não houve casos de erros graves, como ocorreu ao testarmos o método com blocos

menores. Os erros leves podem ser vistos na Figura 17, como os pequenos saltos no erro de estimação ao longo dos quadros, enquanto os erros médios ou discrepantes são os grandes saltos na curva. Os erros graves não podem ser vistos por não existirem para este tamanho de bloco, mas seriam representados por uma sucessão de saltos de altitudes variadas. A Tabela 16 apresentadas os resultados obtidos para o bloco de 48x48 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
48x48	2,84%	829,75	2	2

Tabela 16. Resultados médios obtidos para o bloco único móvel de tamanho 48x48

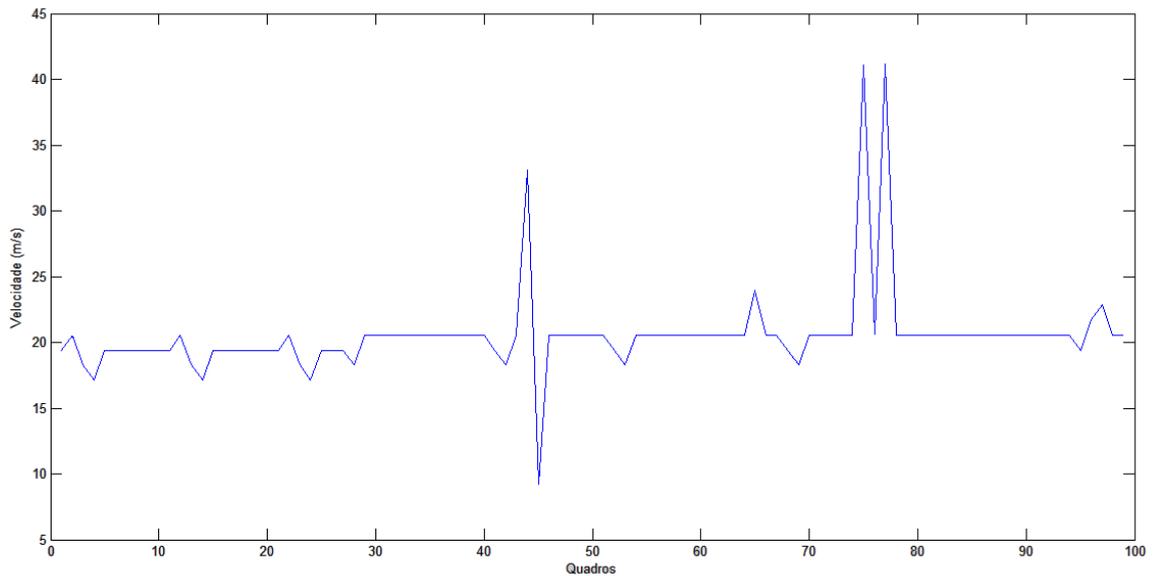


Figura 17. Gráfico dos Erros obtidos para bloco único dinâmico de tamanho 48x48

4.4.2 Bloco de 64x64 pixels

O erro de estimação com blocos de 64x64 *pixels* de tamanho foi de $e=3,54\%$ sem nenhum erro discrepante e sem erros graves. A Figura 18 mostra o comportamento do erro de

estimação da velocidade do trem quadro a quadro, nela podem ser vistos apenas os erros leves que são as pequenas variações no erro. A tabela 17 mostra os resultados obtidos para o bloco de tamanho 64x64 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
48x48	2,84%	829,75	2	2
64x64	3,54%	809,44	0	0

Tabela 17. Resultados médios obtidos para o bloco único estático de tamanho 64x64

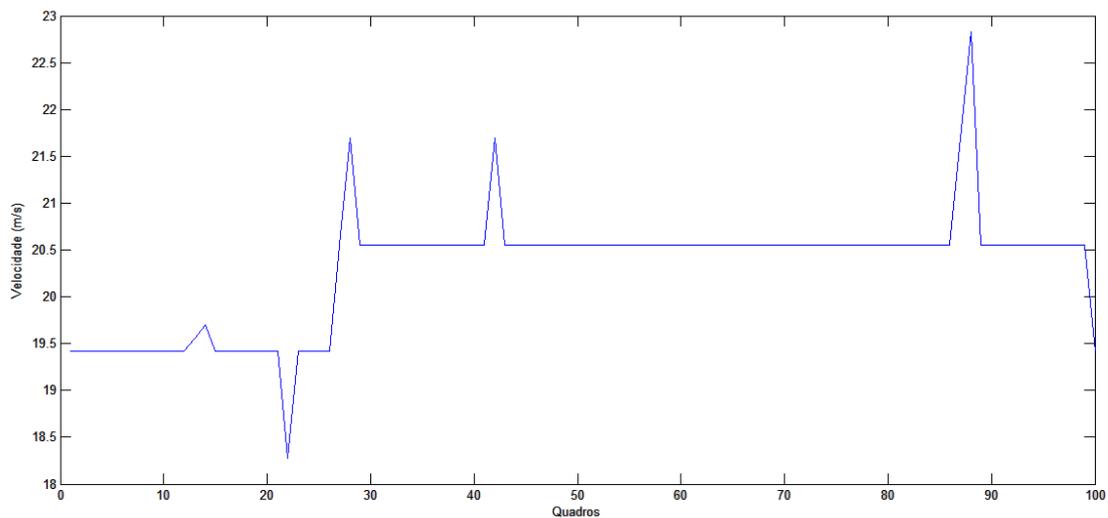


Figura 18. Gráfico dos erros obtidos para bloco único dinâmico de tamanho 64x64

4.5 Dois Blocos Móveis

Desta vez o rastreamento é feito a partir de dois blocos posicionados de acordo com a seção 4.3. e com comportamento descrito em 4.4. Ou seja, dois blocos são rastreados enquanto movem-se ao longo da janela. Neste caso, o tempo de permanência do bloco na janela é dado pelo menor tempo de permanência entre os dois blocos, de modo que quando

um dos dois for reposicionado na origem o outro também é. Novamente, tem-se o dobro de chances de ocorrer erros discrepantes, pois avaliamos dois blocos.

Nesse método percebe-se que a velocidade por ser fruto de cálculos distintos de velocidade, sofrendo, assim, uma suavização dos seus valores mais extremos. Uma vez que o comportamento de um bloco independe do outro, quando estes não produzem valores próximos, que são as estimativas das velocidades dos trens, um valor errado que venha a surgir será parcialmente corrigido. A Tabela 18 apresenta os tempos obtidos com este método.

Tamanho do Bloco	Tempo de Processamento (seg)
8x8	310,24
16x16	299,69
32x32	285,56
48x48	1.994,17
64x64	1.865,76

Tabela 18. Tempos de processamento obtidos para dois blocos móveis com diferentes tamanhos

4.5.1 Bloco de 48x48 pixels

Na configuração com 48x48 *pixels* obteve-se um erro de estimação de $e=2.08\%$ onde devemos levar em conta uma taxa de erros discrepantes de 1,5%. Os resultados obtidos com esse método são os mesmos obtidos com dois blocos fixos de mesmo tamanho. Ficando o tempo de processamento como a única diferença entre os métodos analisados. Neste caso o tempo de processamento foi de 1994 segundos, isto é equivalente a um incremento no tempo de processamento de mais 100%.

A tabela 19 mostra os resultados obtidos para o bloco de tamanho 48x48 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
48x48	2,08%	1.994,17	1	2

Tabela 19. Resultados médios obtidos para dois blocos móveis de tamanho 48x48

4.5.2 Bloco de 64x64 pixels

O resultado obtido com blocos de 64x64 foi o de erro de estimação de $e=3,87\%$, o que confirma um comportamento muito semelhante entre esse método e o método de dois blocos fixos. Novamente a diferença entre eles ficou no tempo de processamento onde o programa com blocos móveis levou 1866 segundos, o que é quase o dobro do seu equivalente com blocos fixos. A tabela 20 mostra os resultados obtidos para o bloco de tamanho 64x64 pixels.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
48x48	2,08%	1.994,17	1	2
64x64	3,87%	1.865,76	0	0

Tabela 20. Resultados médios obtidos para dois blocos móveis de tamanho 64x64

4.6 **Bloco Médio Único**

Este método foi explicado nas seções 3.6 e 3.7.3, e tem por base o método dos blocos móveis. A diferença entre eles está no fato do método de rastreamento não ter mais como base um bloco de referência determinado no quadro anterior, e sim um bloco de referência baseado

na média dos blocos de referência anteriores. Sendo assim, quando a detecção aponta para blocos semelhantes ao bloco que deveria ser encontrado as diferenças são diluídas o que aumenta a robustez do método. Do mesmo modo erros graves são suavizados e, portanto podem ter sua recorrência impedida. O problema deste método está na posição em que se encontra o erro, pois quanto mais afastado este estiver do início do cálculo da velocidade menor será sua influência. Porém se este erro estiver logo no início do cálculo, ele pode levar o algoritmo a perder sua capacidade de detectar o bloco correto. A Tabela 21 apresenta os tempos de processamento obtidos com este método.

Tamanho do Bloco	Tempo de Processamento (seg)
8x8	712,02
16x16	688,30
32x32	654,55
48x48	860,73
64x64	821,99

Tabela 21. Tempos de processamento obtidos para o bloco médio único com diferentes tamanhos

4.6.1 Bloco de 48x48 pixels

Obteve-se com esse bloco um erro de estimativa de velocidade de $e=0,57\%$ e erro de discrepância de 6%. Este último contribui fortemente para a aproximação das velocidades calculadas a velocidade de referência, o que torna inútil a velocidade estimada com bloco de 48x48. A tabela 22 mostra os resultados obtidos para o bloco de tamanho 48x48 pixels.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
48x48	0,57%	860,73	2	4

Tabela 22. Resultados médios obtidos para o bloco médio único de tamanho 48x48

4.6.2 Bloco de 64x64 pixels

Para o bloco de 64x64 *pixels* o erro de estimação ficou em $e=3,63\%$ sem haver nenhuma velocidade discrepante. O valor aqui obtido é pouco maior do que os erros obtidos para blocos fixos e blocos móveis, o que indica que o algoritmo é bem preciso e que por isso acumular o erro ao fazer o bloco de referência com as médias traz mais malefícios do que benefícios. A Tabela 23 apresenta os valores obtidos para blocos de 64x64 *pixel*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
48x48	0,57%	860,73	2	4
64x64	3,63%	821,99	0	0

Tabela 23. Resultados médios obtidos para bloco médio único de tamanho 64x64

4.7 **Dois Blocos Médios**

Esse método adiciona um segundo bloco ao método da subseção 4.6. adicionando assim as vantagens e desvantagens do uso de blocos extras descritos na seção 4.3. A diferença entre o tempo de processamento desse método e do método da subseção 4.5 é pequeno.

Este resultado era esperado, uma vez que a diferença entre os dois métodos é basicamente, a implementação do cálculo da média ponderada na definição das intensidades dos pixels do bloco de referência.

A Tabela 24 apresenta os tempos de processamento obtidos com este método.

Tamanho do Bloco	Tempo de Processamento (seg)
8x8	307,63
16x16	297,17
32x32	283,16
48x48	2.003,50
64x64	1.888,51

Tabela 24. Tempos de processamento obtidos para dois blocos médios com diferentes tamanhos

4.7.1 Bloco de 48x48 pixels

Com dois blocos de 48x48 *pixels* o erro de estimação da velocidade ficou, pela terceira vez, em $e=2,08\%$ e novamente obteve-se 1,5% de taxa de erros discrepantes. Esse resultado reforça a idéia da existência de diferenças de comportamentos na estimação da velocidade de acordo com a escolha do bloco e aponta para a existência de um possível bloco ideal para o cálculo desta velocidade. A tabela 25 mostra os resultados obtidos para o bloco de tamanho 48x48 *pixels*.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
48x48	2,08%	2.003,50	1	2

Tabela 25. Resultados médios obtidos para dois blocos médios de tamanho 48x48

4.7.2 Bloco de 64x64 pixels

Nesta configuração o resultado do erro de estimação foi de 3,87% sem haver valores maiores ou menores em 50% do valor da velocidade de referência. A Tabela 26 apresenta os valores obtidos para os testes realizados com blocos de 64x64.

Tamanho do Bloco	Erro	Tempo de Processamento (seg)	Quadros com velocidade inferior a esperada em 50%	Quadros com velocidade superior a esperada em 50%
48x48	2,08%	2.003,50	1	2
64x64	3,87%	1.888,51	0	0

Tabela 26. Resultados médios obtidos para dois blocos médios de tamanho 64x64

Temos então que os métodos que utilizam blocos de 48x48 pixels, por apresentarem em todos os métodos, resultados errados que influenciam o cálculo da velocidade do trem, não são adequados. Os de 64x64 pixels, em contrapartida, são confiáveis e possuem de modo geral um processamento mais rápido do que os métodos que utilizam blocos menores. Os métodos que se utilizam um bloco apenas tendem a apresentar respostas variadas enquanto os métodos que utilizam dois blocos apresentam comportamento homogêneo.

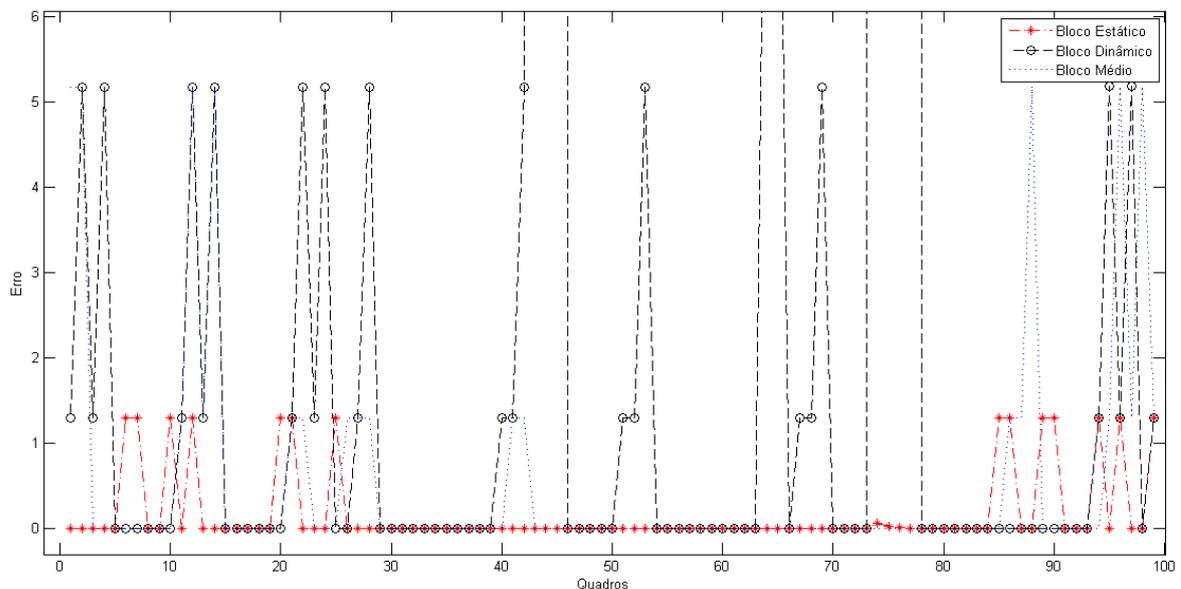


Figura 19. Gráfico dos erros com único bloco para os diferentes tipos de algoritmo

A figura 19 apresenta o gráfico com o comportamento do erro quadro a quadro para os algoritmos que apresentam apenas um único bloco para análise, a partir deles é possível constatar que utilizando apenas um bloco o algoritmo que apresentou melhores resultados foi aqueles que utilizou o bloco estático, isto porque apesar destes apresentar alguns erros pontuais no início e no final do vídeo, seus resultados no restante dos quadros foi igual a zero. O algoritmo de bloco dinâmico foi o que apresentou erros mais significativos e em maior quantidade, devido principalmente aos três tipos de erros comentados na seção 4.4.

A Figura 20 apresenta o comportamento do erro obtido para os diferentes algoritmos testados utilizando dois blocos para estimar o deslocamento, nele é possível perceber que os resultados obtidos utilizando dois blocos foram exatamente iguais. Diferentemente, dos resultados obtidos com um bloco os resultados aqui obtidos apresentam maior quantidade de erro, porém com ordem de grandeza menor quando comparados ao método com um bloco, que leva a concluir que o método com melhores resultados foi aquele que utilizou apenas um bloco para a comparação e com posicionamento fixo ao longo de toda a comparação realizada entre os quadros do vídeo.

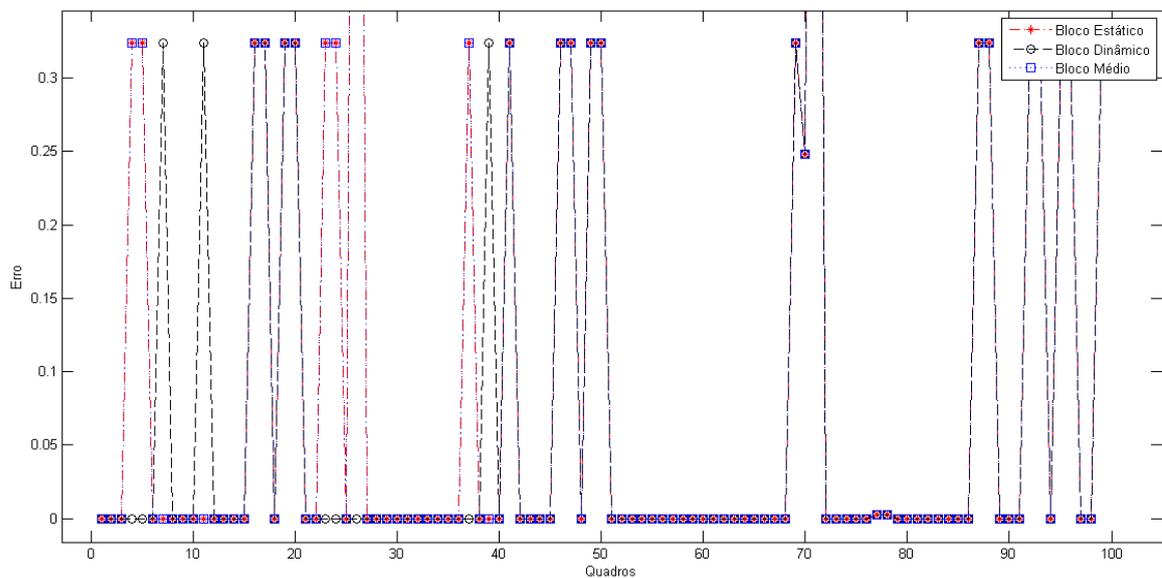


Figura 20. Gráfico dos erros com dois blocos para os diferentes tipos de algoritmo

4.8 Comparação dos resultados obtidos para diferentes direções e distâncias

O local escolhido para a captura dos vídeos permitiu a passagem de trens por diferentes vias alterando, com isso, não só as distâncias entre o dispositivo e o objeto como também o sentido do deslocamento do mesmo.

Conforme descrito na secção 3.6, o posicionamento do bloco de referência em todos os algoritmos analisados é de fundamental importância para que estes alcançassem bons resultados, pois dependendo do posicionamento do bloco de referência o algoritmo poderia se perder. Optou-se por isso pela utilização dos blocos, tanto nos casos onde a comparação foi realizada com apenas um bloco quanto naquelas onde se utilizou com dois, por posicioná-los no centro da janela, com isso o sentido de deslocamento do trem seria observado pelos sinais obtidos.

Ao comparar os resultados alcançados para os vídeos onde o trem desloca-se da esquerda para a direita e aqueles com os trens deslocando-se da direita para a esquerda foi possível perceber que os melhores resultados estavam entre aqueles com deslocamento da esquerda para a direita, conforme mostra a tabela 27. O que levaria a conclusão de que os algoritmos apresentavam falha em um determinado sentido de deslocamento, entretanto analisando as rotas operadas pelos trens em cada uma das vias, constatou-se que os trens que passavam no sentido dos melhores resultados eram trens expressos, ou seja, aqueles que param em poucas estações e por isso apresentam velocidades constantes ao longo de quase toda a via. Nestes casos o erro na estimação da velocidade foi muito pequeno em relação àqueles onde o trem estava saindo ou chegando à estação mais próxima do local de captura dos vídeos.

Sentido	Erro
D-E	4,03%
E-D	2,63%
INCLINADO	4,19%

Tabela 27. Erros obtidos para os trens em diferentes direções e com inclinação

Via	Distância Dispositivo - Trem (m)	Erro
1	13,5	3,87%
2	17,2	19,85%
3	20,9	2,05%
4	24,6	4,15%

Tabela 28. Erro obtido para diferentes distâncias

Os resultados alcançados para trens em aclave foram de 4%, porém neste caso tem-se o mesmo problema acima descrito, o local de captura dos vídeos para esse caso era próximo de uma estação, o acabou por causar o aumento no erro dos resultados encontrados.

Esta suposição pode ser confirmada quando os vídeos foram agrupados e analisados de acordo com a via de passagem. A Tabela 28 apresenta os resultados dos erros encontrados para cada uma delas.

As velocidades obtidas pelos algoritmos diferiram muito quando o trem estava em aceleração, vias 1 e 2 da Tabela 28, viu-se que para estes casos o algoritmo é mais preciso, pois calcula a velocidade instantânea enquanto o sistema de controle calcula a velocidade média ao longo do circuito de via. Nos casos onde o trem percorre o trecho em aceleração a velocidade calculada pelo sistema não pode ser utilizada para comparação, pois sempre apresentará um valor muito diferente daquele obtido pelos algoritmos utilizados.

As vias 3 e 4 referem-se aos trens com velocidades mais constantes ao longo da via devido as características de percursos a eles imposta. Os erros nestas vias foram menores, devido, principalmente, a rota traçada pelo CCO para tais trens, neste caso a velocidade instantânea se aproximou da velocidade com que o trem percorreu o trecho.

A fim de ratificar tal teoria, novas medições foram realizadas, porém nesta nova abordagem foram comparados os valores obtidos através do sistema, àqueles obtidos com este trabalho e a velocidade aferida no velocímetro do trem. Em média, os valores obtidos para os trechos com aceleração utilizando o algoritmo proposto apresentaram erros muito menores quando comparados com os erros obtidos com a velocidade obtida pelo circuito de via. A Tabela 29 apresenta os resultados obtidos com os sistemas. O sistema vigente nela mencionado faz referência ao sistema utilizado pela operadora do sistema de trens, enquanto o sistema proposto refere-se ao algoritmo desenvolvido neste trabalho.

Sistema	Erro
Vigente	1,28%
Proposto	55,92%

Tabela 29. Erro na estimação da velocidade para os sistemas analisados

CONCLUSÃO

Neste trabalho foi proposto um sistema para a estimação da velocidade dos veículos ferroviários utilizando processamento de vídeos. O algoritmo utilizou um método comparativo entre quadros adjacentes para estimar o deslocamento realizado. O critério de similaridade utilizado foi o SAD. O sistema mostrou-se bastante robusto no que tange o rastreamento do movimento, visto que os resultados encontrados mostraram que este é capaz de identificar o deslocamento ocorrido de um determinado padrão entre um quadro e o seu sucessor. Além disso, em todos os métodos utilizados o sentido do deslocamento do trem foi classificado corretamente, mostrando que independente do sentido do deslocamento ele é capaz de estimar o deslocamento ocorrido entre os quadros.

O sistema mostrou-se sensível ao deslocamento, visto que todos os quadros analisados já apresentavam parte ou todo o trem em sua área, o que permitiu reduzir o tempo de processamento analisando informações inúteis ao processo.

O algoritmo proposto testou um total de seis diferentes abordagens para avaliação daquela que melhor se enquadrava no sistema em análise. Os resultados obtidos em vias onde a aceleração era próxima de zero pelo algoritmo, pelo sistema em vigor e aqueles aferidos no velocímetro foram muito próximos e por isso neste caso os dois sistemas mostraram-se úteis. Nos casos onde os trechos analisados estavam em aceleração a velocidade obtida com o algoritmo e a coletada no velocímetro do trem foram muito próximas. Já as comparações realizadas entre as velocidades informadas pelo sistema vigente e aquelas observadas no velocímetro do trem mostraram-se divergentes.

Através desta análise é possível constatar que o sistema proposto apresentou-se como uma solução de baixo custo e bastante robusta no processo, podendo vir a ser utilizada em conjunto com o sistema já em utilização. Vale ressaltar, no entanto, que o sistema hoje utilizado na empresa testada, não têm como função principal estimar a velocidade, sua atividade principal, para o qual foi desenvolvido é a sinalização ferroviária.

Um conjunto de câmeras instaladas ao longo da via, além de permitir a detecção do movimento para estimar a velocidade, permitiria o monitoramento pelos controladores da via. Permitindo que estes, possam por exemplo em caso de acidente avaliar o que está acontecendo no local, permitindo uma tomada de decisão melhor suportada.

Além de permitir o acompanhamento de manutenções preditivas e corretivas, monitoramento de situações adversas, tais como invasão da via e identificação dos motivos da parada de um trem.

Possibilidades futuras

O sistema neste trabalho proposto não permite resultados em tempo real, por isso uma abordagem a ser estudada no futuro seria a aplicabilidade de um sistema como esse em tempo real.

As vantagens do sistema em tempo real seria a de melhor controle da velocidade dando ao controlador informações mais criteriosas para o planejamento de rotas com maior precisão. Esta precisão permitiria que mais trens pudessem circular ao mesmo tempo na via, aumentando a capacidade de atendimento aos usuários, reduzindo assim os gargalos no horários de maior movimento.

REFERÊNCIAS

- [1] ALBUQUERQUE, A. Processamento de Imagens: Métodos e Análises. Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro.
- [2] ALEXANDRE, L. Detecção de movimento e sua aplicação à monitorização do tráfego rodoviário. 1997. 91f. Dissertação (Mestre em Engenharia Electrotécnica e de Computadores) – Faculdade de Ciências, Universidade do Porto, 1997.
- [3] ALLEN, P; TIMCENKO; A. YOSHIMI, B; MICHELMAN, P. Automated tracking and grasping of a moving object with a robotic hand-eye system. IEEE Transactions on Robotics and Automation, v.9, n.2, p.152-165, apr. 1993.
- [4] ALLEN, P; YOSHIMI, B; TIMCENKO, A. Real-time visual servoing. Department of Computer Science - Columbia University, New York, p. 1-12.
- [5] ARAÚJO, A. Desenvolvimento de equipamento para medir a vazão de arrasto em bancos de areia de estuários rasos. Departamento de Engenharia Mecânica, Univerisdade Federal de Pernambuco, Recife.
- [6] Bertoni, M. Dispositivos para ajuste remoto de velocidade em trem com vagão com restrição de velocidade. Negócio nos trilhos – Encontrem 2005, São Paulo, out. 2005.
- [7] CAMPOS-NETO, C. Setor ferroviário brasileiro. Desafios do desenvolvimento, Brasília, n.55, p. 24, out/nov. 2009.
- [8] CARDOSO, E. Análise do transporte coletivo urbano sob a ótica dos riscos e carências sociais. 2008. 38f. Dissertação (Doutorado em Serviço Social) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2008.
- [9] CARVALHO, C. Desenvolvimento urbano nos trilhos. Desafios do desenvolvimento, Brasília, n.55, p. 27, out/nov. 2009.

- [10] CAVALCANTE, A; LEÃO-NETO, C. Sistema de medição de velocidade via webcam. 2006. 35f. Monografia (Bacharelado em Engenharia da Computação) – Instituto de Estudos Superiores da Amazônia, Belém, 2006.
- [11] CHEEZUM, M. WALKER, W; GUILFORD, W. Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophysical Journal*, Charlottesville, v. 81, p. 2378-2388, out 2001.
- [12] CONFEDERAÇÃO NACIONAL DO TRANSPORTE. Boletim Estatístico. Disponível em <<http://www.cnt.org.br>>. Acesso em 12 de Maio de 2011.
- [13] CORDEIRO, C. et al. A visão do cliente para melhoria da qualidade do sistema de transporte coletivo por ônibus. *ENEGEP/ABEPRO*, p. 1592-1599, Porto Alegre, out/nov. 2005.
- [14] CORREIA, M. Análise de movimento em sequência de imagens. 1995. 153f. Tese (Mestrado em Engenharia) - Faculdade de Engenharia da Universidade do Porto, Porto, 1995.
- [15] CUPILLARD, F, AVANZI, A, BREMOND, F. AND THONNAT, M. Video understanding for metro surveillance, p.1-6.
- [16] DRUMMOND, T; CIPOLLA, R. Real-time visual tracking of complex structures. *IEE Transactions on Pattern Analysis and Machine Intelligence*, v. 24, n. 7, p. 932-946, jul. 2002.
- [17] FARIA, E. História dos transportes terrestres no mundo. Universidade Federal do Rio de Janeiro.
- [18] FERNÁNDEZ, Francisco. *Señalización Ferroviaria Del guardagujas a la operación sin conductor*. Espanha: Malpe, 2009.
- [19] FERREIRA, R. Codificador H.264/AVC com compensação de movimento baseada em partições alternativas de macrobloco. 2009. 102f. Dissertação (Mestrado em Engenharia Elétrica) – Faculdade de Tecnologia, Universidade de Brasília, Brasília, 2009.

- [20] GONZALEZ, R., WOODS, R. *Processamento Digital de Imagens*. 3 ed., São Paulo:Pearson Prentice Hall, 2010.
- [21] HUNG, E. Compensação de movimento utilizando blocos multi-escala e forma variável em um codec de vídeo híbrido. 2007. 87f. Dissertação (Mestrado em Engenharia Elétrica) – Faculdade de Tecnologia, Universidade de Brasília, Brasília, 2007.
- [22] LAI, Y. et al. Machine Vision Analysis of the Energy Efficiency of Intermodal Trains. Railroad Engineering Program, University of Illinois at Urbana-Champaign, 2005.
- [23] LIM, J. *Two Dimensional Signal and Image Processing*. Englewood Cliffs, New Jersey, Prentice Hall,1990.
- [24] LOPES-JUNIOR, E. Uso de vídeo e rede de alta velocidade para monitoração em tempo real de abelhas sem ferrão. 2009. 67f. Dissertação (Mestrado em Engenharia) - Escola Politécnica, Universidade de São Paulo, São Paulo 2009.
- [25] MAIA-JUNIOR, L; COSTA, A. Monitoramento em tempo real baseado em fluxo ótico. Escola Politécnica, Universidade de São Paulo, São Paulo.
- [26] MASCARENHAS, Nelson., AND VELASCO, Flávio. *Processamento Digital de Imagens*, 2ª edição, Buenos Aires, Kapelusz,1989.
- [27] MINHONI, D. Detecção de movimento de objetos em tempo real utilizando dispositivos de lógica programável complexa. 2006. 147f. Dissertação (Mestrado em Engenharia) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Paulo, 2006.
- [28] OLIVARES, J. et al. Minimum Sum of Absolute Differences implementation in a single FPGA device. University of Málaga and University of Córdoba, Spain.
- [29] OLIVEIRA, I. Metodologia para detecção rápida de movimento em sequências de imagem. 2003. 191f. Tese (Doutorado em Ciências: Física Aplicada) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2003.

- [30] PINHO, R; TAVARES, J; CORREIA, M. Introdução à Análise de Movimento usando Visão Computacional. 2004. 49p. Relatório Interno - Faculdade de Engenharia, Universidade do Porto, Porto, 2004.
- [31] PORTO, M. Avaliação do algoritmo diamond search para estimação de movimento com múltiplos quadros de referência. Dissertação (Mestrado em Ciência da Computação) – Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.
- [32] PORTO, R. Desenvolvimento arquitetural para estimação de movimento de blocos de tamanhos variáveis segundo o padrão H.264/AVC de compressão de vídeo digital. 2008. 96f. Dissertação (Mestrado em Ciência da Computação) – Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.
- [33] RACHEL, F. Proposta de um controlador automático de trens utilizando lógica nebulosa preditiva. 2006. 162f. Dissertação (Mestrado em Engenharia) - Escola Politécnica, Universidade de São Paulo, São Paulo, 2006.
- [34] RAMPINELLI, M. REINAN, T. AND FERREIRA, E. Análise da robustez da visão computacional de topo no controle de movimentos de robôs móveis ou veículos multi-articulados. Departamento de Engenharia Elétrica, Universidade Federal do Espírito Santo, Espírito Santo.
- [35] REIS-JUNIOR, A. AND SOARES-FILHO, M. Aplicações de processamento de imagens a sistemas de segurança. 2002. 21f. Trabalho (Curso de Ciências da Computação) – Setor de Ciências Exatas, Universidade Federal do Paraná, Paraná, 2002.
- [36] RICHARDSON, I. H.264 and MPEG-4 video compression: video coding for next-generation multimedia. Aberdeen, Wiley Editorial, 2003. 306p.
- [37] ROSA, L et al. Avaliação algorítmica para a estimação de movimento na compressão de vídeos digitais. Hifen, Uruguaiana, v.31, n. 59/60, 2007a.

- [38] ROSA, L. Investigação sobre algoritmos para a estimação de movimento na compressão de vídeos digitais de alta definição: uma análise quantitativa. Monografia (Bacharelado em Ciência da Computação) – Universidade federal de Pelotas, Pelotas, 2007b.
- [39] SILVA, E. Transcodificador de vídeo Wyner-Ziv/H.263 para comunicação entre dispositivos móveis. 2008. 117f. Dissertação (Mestrado em Engenharia Elétrica) – Faculdade de Tecnologia, Universidade de Brasília, Brasília, 2008.
- [40] SUN, K. Adaptive step-size motion estimation based on statistical sum of absolute differences. United States Patent, California, p. 1-10, jan. 2000.
- [41] WANG, Y. et al. Improving truck and speed data using paired video and single-loop sensors. Department of Civil and Environmental, Engineering University of Washington, 2006, p.115. Final Technical Report.
- [42] WONG, S; VASSILIADIS, S; COTOFANA, S. Sum of Absolute Differences Implementation in FPGA Hardware. Electrical Engineering Department, Delft University of Technology.
- [43] YOSHIMURA, H. AND YOSHIKOSHI S. *Railway Signal*. Japan Association of Signal Industries, 1977.