

FABIO SANTOS BERNARDES

**Sistema para detecção de desalinhamento de viga-guia em regiões
de mudança de via dos Sistemas Monotrilho –
*Sistema de visão***

**São Paulo
2016**

FABIO SANTOS BERNARDES

**Sistema para detecção de desalinhamento de viga-guia em regiões
de mudança de via dos Sistemas Monotrilho –
*Sistema de visão***

Monografia apresentada à Escola
Politécnica da Universidade de São
Paulo para obtenção do título de
Especialista em Tecnologia
Metroferroviária.

**São Paulo
2016**

FOLHA DE APROVAÇÃO

FÁBIO SANTOS BERNARDES

**Monografia apresentada à Escola
Politécnica da Universidade de
São Paulo para obtenção do título
de Especialista em Tecnologia
Metro-Ferroviária.**

**Área de Concentração:
Engenharia de Computação**

Aprovado em:

Banca Examinadora

Prof. Dra. Anna Helena Reali Costa

Universidade de São Paulo

Assinatura:

Prof. Dr. Jorge Rady de Almeida Junior

Universidade de São Paulo

Assinatura:

Prof. Dr. Paulo Sérgio Cugnasca

Universidade de São Paulo

Assinatura:

FABIO SANTOS BERNARDES

**Sistema para detecção de desalinhamento de viga-guia em regiões
de mudança de via dos Sistemas Monotrilho –
*Sistema de visão***

Monografia apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Especialista em Tecnologia Metroferroviária.

**Área de Concentração:
Tecnologia Metroferroviária**

**Orientador: Prof. Dra.
Anna Helena Reali Costa**

**São Paulo
2016**

Catalogação-na-publicação

Bernardes, Fábio Santos

Sistema para detecção de desalinhamento de viga-guia em regiões de
mudança de via dos Sistemas Monotrilho - Sistema de Visão / F. S. Bernardes
- São Paulo, 2016.

37 p.

Monografia (Especialização em Tecnologia Metroferroviária) - Escola
Politécnica da Universidade de São Paulo. PECE – Programa de Educação
Continuada em Engenharia.

1. Sistema de Visão I. Universidade de São Paulo. Escola Politécnica.
PECE – Programa de Educação Continuada em Engenharia II.t.

AGRADECIMENTOS

Aos meus familiares pela compreensão e acolhimento.

As gerências da Companhia do Metropolitano de São Paulo e aos Engenheiros Agnelo Gomes, Carlos Paixão e Paulo Beu, pela demonstração de confiança na competência profissional ao me premiar com a eleição para o curso de especialização em Tecnologia Metroferroviária, do mesmo modo ao apoio incondicional concedido para o cumprimento desta trajetória com longanimidade.

À Doutora Anna Reali que com gáudio aceitou e sabiamente aconselhou seus orientados.

Ao Engenheiro Ivan Maestrelli pelo auxílio para obtenção das informações relacionadas ao material rodante.

“O passado é lição para refletir, não para repetir.”

RESUMO

Um dos principais riscos operacionais durante a operação de um sistema metroferroviário é a condução de um trem em modalidade manual através de uma região de mudança de via em condição de desalinhamento de rota, situação favorável para a ocorrência do descarrilamento da composição. Este risco é potencializado no modal de transporte Monotrilho, onde as regiões de mudança de via são elevadas – 15 metros de altura em média – e as consequências de um acidente desta natureza podem representar grandes prejuízos ambientais, patrimoniais e civis. Com o propósito de mitigar esse risco, a proposta de um sistema de processamento e reconhecimento das condições de desalinhamento de via, chamado de SDDV – Sistema de Detecção de Desalinhamento de Via – possui, como parte integrante, o sistema de câmeras embarcadas do trem, que coletam as imagens da via para, então, processá-las e identificar a descontinuidade das vigas através de algoritmos. Como parte integrante deste sistema, o objetivo deste trabalho é apresentar, de forma mais detalhada, as principais características do sistema de câmeras embarcadas do trem, as avaliações das capturas das imagens e propostas de relação entre distâncias reais e distâncias em pixels da imagem capturada.

ABSTRACT

A major operational risk during the operation of a subway-railroad system is driving a train in manual mode by a turnout region in a misalignment condition, favorable situation for the derailment of the composition. This risk is enhanced in the monorail transport modal, where track-switches regions are high – 15 meters high on average – and the consequences of an accident of this nature can pose major environmental, property and civil damages. In order to mitigate this risk, the proposal of a processing system and recognition of track misalignment conditions, called SDDV – Track Misalignment Detection System – has, as part of the embedded train cameras system, that collect the images of the track to process it and identify the discontinuity of the beans through algorithms. As part of this system, the aim of this study is to present in more details the main characteristics of the embedded train cameras system, evaluations of the captured images and proposed relationship between real distances and distances in pixels of the captured image.

LISTAS DE FIGURAS

FIGURA 1.1 – TRAÇADO DA LINHA 15-PRATA DO MONOTRILHO E SUAS ESTAÇÕES EM DESTAQUE.....	15
FIGURA 1.2 – UM APARELHO DE MUDANÇA DE VIA – AMV – DO MONOTRILHO DE LAS VEGAS (EUA).....	16
FIGURA 2.1 – ESPECTRO ELETROMAGNÉTICO COM A AMPLIAÇÃO DO ESPECTRO VISUAL DA LUZ.....	20
FIGURA 2.2 – CÍRCULOS CROMÁTICOS DOS SISTEMAS DE CORES RGB E CMY, RESPECTIVAMENTE.....	22
FIGURA 2.3 – PROCESSO DE SEPARAÇÃO DAS CORES ATRAVÉS DO “FILTRO DE BAYER”; A LUZ É FILTRADA E CADA UMA DAS TRÊS CORES (RGB) INCIDEM SOBRE SENSORES ESPECÍFICOS PARA CADA COR.	23
FIGURA 3.1 – DIAGRAMA DE INTERLIGAÇÃO DO SISTEMA DE MONITORAMENTO DO TREM. EM DESTAQUE A CONFIGURAÇÃO DE UM CARRO DE EXTREMIDADE (A1) E O CARRO INTERMEDIÁRIO (C2), ONDE ESTÃO LOCALIZADAS AS VCCU'S. FONTE: ADAPTADO DA DOCUMENTAÇÃO DO PROJETO MONOTRILHO LINHA 15-PRATA.	24
FIGURA 3.2 – ILUSTRAÇÃO INDICANDO A LOCALIZAÇÃO DA CÂMERA VOLTADA PARA O EXTERIOR DO TREM E À DIREITA O SEU DETALHE. NESTA CONFIGURAÇÃO AS IMAGENS CAPTURADAS PODEM SER UTILIZADAS PARA MONITORAMENTO DAS CONDIÇÕES À FREnte DA VIA, LOCAL ONDE EXISTE A MOVIMENTAÇÃO DO TREM.	25
FIGURA 3.3 – IMAGEM EXTRAÍDA DE UM QUADRO DE VÍDEO CAPTURADO PELA CÂMERA EXTERNA DO TREM. ESTA IMAGEM FOI EXTRAÍDA DE UM VÍDEO COM RESOLUÇÃO DE 800 X 600 COM TAXA DE 12 FPS - FRAMES PER SECOND.....	26
FIGURA 3.4 – IMAGEM DA VIA NORMAL, SEM NENHUMA INTERFERÊNCIA EXTERNA.	28
FIGURA 3.5 – IMAGEM DA VIA COM A PRESENÇA DE SOMBRA PROVOCADA PELAS EDIFICAÇÕES DOS ARREDORES.	28
FIGURA 3.6 – IMAGEM DA VIA COM A PRESENÇA DE SOMBRA PROVOCADA PELA COBERTURA DA ESTAÇÃO VILA PRUDENTE (LINHA 15-PRATA).....	28
FIGURA 3.7 – IMAGEM DA VIA COM CURVA SUAVE À ESQUERDA.....	28
FIGURA 3.8 – IMAGEM DA VIGA-GUIA METÁLICA DO AMV (NOTAR A DIFERENÇA DE COR E SOMBA DE ESTRUTURAS VIZINHAS).	28
FIGURA 3.9 – IMAGEM DA VIA COM A PRESENÇA DE SOMBRA PROVOCADA PELA COBERTURA DA ESTAÇÃO ORATÓRIO (LINHA 15-PRATA)	28
FIGURA 3.10 – IMAGEM DA VIA COM A PRESENÇA DE SOMBRA PROVOCADA POR UM TREM EM DIREÇÃO OPosta, NA VIA AO LADO.....	29
FIGURA 3.11 – VISTA DO TREM SITUADO NA OFICINA DE MANUTENÇÃO, ILUMINADA, COM VISÃO PARA ÁREA EXTERNA DURANTE A NOITE.	29
FIGURA 3.12 – IMAGEM DA VIA DENTRO DO PATIO DE MANUTENÇÃO DURANTE A NOITE E SEM ILUMINAÇÃO.	29
FIGURA 3.13 – IMAGEM DA VIA DENTRO DO PATIO DE MANUTENÇÃO DURANTE A NOITE, SEM ILUMINAÇÃO COM CURVA SUAVE À DIREITA.	29

FIGURA 3.14 – IMAGEM DA VIA DURANTE A NOITE E SEM ILUMINAÇÃO, COM A LENTE MOLHADA PELA CHUVA.....	29
FIGURA 3.15 – IMAGEM DA VIA DURANTE A NOITE, SEM ILUMINAÇÃO, COM A PRESENÇA DAS LUZES DE SINALIZAÇÃO DO AMV.....	29
FIGURA 3.16 – IMAGENS EXTRAÍDAS DO VÍDEO COM O TREM PARADO A (A) 12,0 METROS E A (B) 2,8 METROS DA PLACA DE FIM DE VIA, MARCAÇÃO 4 DA IMAGEM.....	31
FIGURA 3.17 – A LINHA EM VERDE REPRESENTA A QUANTIDADE DE PIXELS VERTICais DA BASE DA IMAGEM ATÉ A BASE DO POSTE DA PLACA DE FIM DE VIA. OS QUADROS À ESQUERDA E À DIREITA SÃO RESPECTIVAMENTE AS AMPLIAÇÕES DAS IMAGENS CAPTURADAS A 12,0 METROS E 2,8 METROS DA PLACA.....	32
FIGURA 3.18 – MODELO INCORPORANDO A ALTURA DA CÂMERA NA DETERMINAÇÃO DA RELAÇÃO PIXEL POR METRO DO SISTEMA. NA REPRESENTAÇÃO SÃO CONSIDERADOS OS ÂNGULOS DE PROJEÇÃO EM RELAÇÃO AO PLANO DA VIA NOS CENÁRIOS ONDE A DISTÂNCIA AO PONTO DE REFERÊNCIA, OU PLACA DE FIM DE VIA, É (A) 12,0 METROS E (B) 2,80 METROS.....	33

LISTAS DE TABELAS

TABELA 3.1 – CARACTERÍSTICAS TÉCNICAS DAS CÂMERAS EXTERNAS DO TREM. 27

LISTA DE SIGLAS E ABREVIATURAS

AMV	Aparelho de Mudança de Via
AUTO	Modo de Operação Automático
AVI	<i>Audio Video Interleave</i>
CCD	<i>Charge-Couple Device</i>
CMOS	<i>Complementary Metal Oxide Semiconductor</i>
CMY	<i>Cyan, Magenta and Yellow</i>
FPS	<i>Frames per Second</i>
FTP	<i>File Transfer Protocol</i>
GPS	<i>Global Positioning System</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
IRSE	<i>Institution of Railway Signal Engineers</i>
MAN	Modo de Operação Manual
MCMT	Modo de Controle Manual do Trem
MCS	Modo de Controle Manual Supervisionado
METRO	Companhia do Metropolitano de São Paulo
PIXEL	<i>Picture Element</i>
RFID	<i>Radio Frequency Identification</i>
RGB	<i>Red, Green and Blue</i>
SDDV	Sistema de Detecção de Desalinhamento de Via
USB	<i>Universal Serial Bus</i>
VCCU	<i>Vehicle Communications Unit</i>
VDVR	<i>Vehicle Digital Video Recorder</i>

SUMÁRIO

1.	INTRODUÇÃO.....	15
1.1.	PROPOSTA DO TRABALHO	17
1.2.	USO DE IMAGENS EM SISTEMAS METROFERROVIÁRIOS.....	18
1.3.	ORGANIZAÇÃO DO TRABALHO	19
2.	SISTEMA DE VISÃO.....	20
2.1.	ESTRUTURA DO OLHO HUMANO	20
2.2.	SISTEMA DE CORES.....	21
2.3.	MODELO RGB	22
2.4.	IMAGEM DIGITAL.....	22
3.	DESENVOLVIMENTO.....	24
3.1.	CÂMERA DO TREM	24
3.2.	RELAÇÃO ENTRE A DISTÂNCIA REAL E A DISTÂNCIA EM PIXELS.....	30
4.	CONCLUSÃO.....	35
5.	REFERÊNCIAS BIBLIOGRÁFICAS.....	37

1. INTRODUÇÃO

Para cumprir o compromisso de atender ao direito dos cidadãos de consumar os seus deslocamentos, a Companhia do Metropolitano de São Paulo – METRÔ apresentou as novas linhas 15-Prata e 17-Ouro, com tecnologia em monotrilho, projetadas para o atendimento das demandas de transporte da região sul, sudeste e de parte da região leste da cidade de São Paulo. Contando com uma estrutura elevada, com altura média equivalente a 15 metros, o monotrilho é um sistema de trem sobre pneus, com tração elétrica e que se desloca unido a vigas de concreto que, apoiadas sobre pilares, constituem os traçados de percurso, com extensões próximas a 22 km, entre as principais avenidas e vias secundárias destas regiões, conforme representado na Figura 1.1. Uma forte característica em regiões centrais de grandes cidades é a existência de uma maior interferência visual deste tipo de obra sobretudo em trechos onde são exigidas a incorporação das regiões de mudanças de via (Kouassi, 2001).

As regiões de mudança de via são regiões em que as vias se cruzam através de vigas móveis, no caso do monotrilho, mantendo o apoio e continuidade física necessários para orientação da passagem suave das rodas durante a mudança da via, onde encontra-se, inicialmente, o trem em curso (IRSE, 2011). Nas linhas do monotrilho, o conjunto formado pelas estruturas metálicas e dispositivos elétricos dentro de uma região de mudança de via é chamado aparelho de mudança de via – AMV, conforme Figura 1.2.

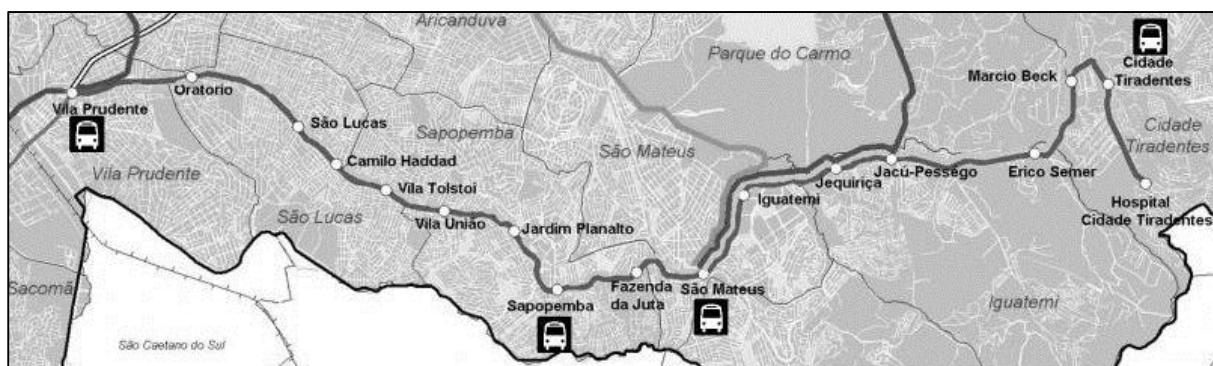


Figura 1.1 – Traçado da Linha 15-Prata do monotrilho e suas estações em destaque.



Figura 1.2 – Um aparelho de mudança de via – AMV – do monotrilho de Las Vegas (EUA).

Nesta imagem a viga metálica do AMV de substituição, localizada na região central da imagem, encontra-se posicionada de modo a favorecer a mudança de trajeto do trem entre as vias.

Os trens do monotrilho são projetados para oferecem três modos de operação configuráveis, denominados: modo de operação automática do trem (AUTO), modo de operação com controle manual supervisionado (MCS) e modo de operação manual (MAN). Os sistemas de controle de movimentação automática de trens são projetados para atingirem baixas taxas de falhas e reagirem de maneira segura quando um evento inesperado ocorrer. Porém, o fato do sistema se manter em condição segura durante a condição de falha não representa que seus serviços ou funções continuarão disponíveis. Assim, nestas situações, a ausência das funcionalidades providas pelo sistema deverá ser coberta por agentes externos que assumirão o papel de analisar e atuar de forma segura durante as operações executadas a partir do evento de falha.

O exemplo claro onde pode ocorrer esta situação, considerado uma das principais condições de riscos de um trem no sistema monotrilho, é a sua condução durante mudanças de via em condição de desalinhamento da viga metálica. Este risco é potencializado nos monotrilhos, onde as regiões de AMV são elevadas (em

média 15 metros de altura) e as consequências de um acidente desta natureza representam grandes prejuízos ambientais, patrimoniais e civis.

Sistemas com intuito de mitigar esse risco, também conhecidos como sistemas auxiliares de via (IEEE ,2004), são conhecidos e divididos em duas grandes categorias: os sistemas *trip-stop*, que baseiam-se em atuadores mecânicos ou magnéticos que atuam diretamente sobre o sistema pneumático do trem quando se desloca sobre esse dispositivo na via, acionando fisicamente o freio de emergência; e os sistemas baseados em sensoriamento da via, que são acionados em função das informações providas pelos sensores de posicionamento da via, que comunicam com o módulo de controle manual embarcado, provocando o disparo de um comando elétrico de frenagem de emergência. As principais características destes sistemas são: custo alto, arquitetura complexa e nível de manutenção médio.

Uma proposta alternativa é através de um sistema de processamento e reconhecimento das condições de desalinhamento de via, chamado SDDV (Sistema de Detecção de Desalinhamento de Via). Sua principal característica é utilizar sequências de vídeo captadas por uma câmera embarcada no trem do monotrilho e processá-las, em tempo real, para que gerem resultados que permitam informar ao módulo do trem a necessidade de aplicação do freio de emergência. Essa frenagem acontece somente caso haja viga metálica do AMV na posição de desalinhamento de rota, evitando-se assim prováveis acidentes.

1.1. Proposta do trabalho

O objetivo deste trabalho é apresentar, como parte integrante do Sistema de Detecção de Desalinhamento de Via – SDDV – os estudos detalhados do sistema de visão das câmeras embarcadas no trem e as características da captura de imagem. Desta forma, integra e complementa o estudo do sistema de processamento e reconhecimento das condições de desalinhamento de via.

É importante ressaltar que, muito embora estudos específicos desta aplicação tenham sido apresentados no ambiente metroferroviário, não foram localizadas citações a trabalhos equivalentes direcionados à tecnologia monotrilho, o que reitera a relevância em se investigar a validade do objetivo proposto por meio deste trabalho.

1.2. Uso de Imagens em Sistemas Metroferroviários

Sistemas auxiliares que adotam métodos de aquisição, processamento, análise e interpretação de imagens são bastante aplicados no auxílio e complementação dos sistemas de proteção automáticos ligados aos sistemas metroferroviários.

Lai et al. (2007) desenvolveram um sistema de escaneamento de trens de carga em movimento a partir de técnicas de processamento de imagens para captura e estimação de parâmetros aerodinâmicos. Estes parâmetros foram utilizados para estimar, a partir de outras imagens, o carregamento de vagões, possibilitando a obtenção de um modelo aproximado do consumo energético em função da distribuição de cargas e uma alternativa para a prevenção de acidentes devido a sobrecargas.

A partir da captura de imagens dos aspectos de sinaleiros – dispositivo de indicação visual que exibe instruções ou fornece aviso prévio de instruções, baseadas em cores, relativas à autorização para o condutor de um trem prosseguir viagem (IRSE, 2011) – Losarkul e Suthakorn aplicaram pré-processamento digital e técnicas de redes neurais para obterem o reconhecimento de maneira autônoma dos sinais de tráfego. Tal solução, aplicada em situações de procedimentos operacionais, onde manobras com operadores a bordo se façam necessárias e o sistema de proteção automático esteja desligado ou inativo, permite a atuação deste sistema no trem como ação mitigadora em reação a atos inesperados ou inadvertidos por parte do operador diante de uma indicação restritiva à movimentação da composição.

Quando a comunicação via rádio é adotada para o sistema de controles de trens, a avaliação da qualidade deste canal torna-se importante. Na metodologia de modelagem de via proposta por Hattori e Kato (2004) a utilização de um trem piloto, equipado com escâner laser, câmeras para composição de informação estéreo do ambiente e sistemas de posicionamento baseado na tecnologia GPS, permitiu a obtenção de um modelo de propagação tridimensional para avaliação das variações das características ambientais que influenciam na qualidade do canal de comunicação do sistema.

1.3. Organização do trabalho

Esta monografia está estruturada da seguinte forma:

O capítulo 2 apresenta uma visão geral do Sistema de Visão, abordando de forma geral desde a formação da imagem no olho humano, até a representação nos sistemas de cores e formação da imagem digital.

No capítulo 3 estão apresentadas, de forma detalhada, as características da câmera embarcada do trem, a análise de suas principais características e proposta para as relações entre distâncias reais e distâncias em pixels da imagem.

No capítulo 4 são apresentadas as conclusões e as propostas de trabalhos futuros.

No capítulo 5 são apresentadas as referências bibliográficas utilizadas para a elaboração desta monografia.

2. SISTEMA DE VISÃO

2.1. Estrutura do Olho Humano

O processo de formação de uma imagem pelo globo ocular só é possível com a emissão de uma fonte de luz, seja pelo sol, lâmpada ou fogo. Entende-se a natureza da luz viajando no espaço como uma dualidade onda-partícula, em que possui características simultâneas de partícula, no transporte de fótons, e de onda, no transporte de radiação eletromagnética. Como onda, ela compõe parte do espectro da radiação eletromagnética visível ao olho humano. Conforme Figura 2.1 a seguir, com a variação da frequência da onda (ou de seu comprimento de onda) existe uma variação na cor da luz.

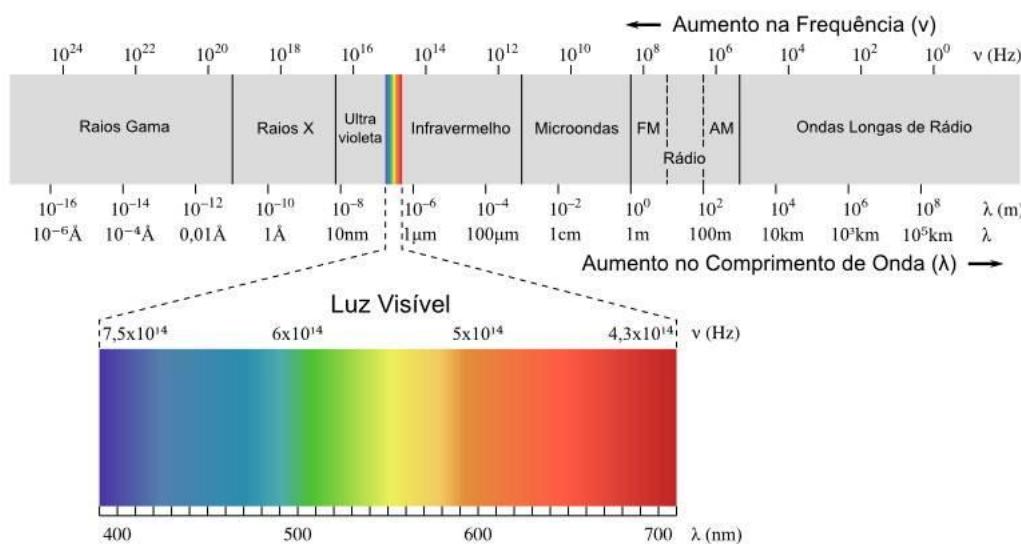


Figura 2.1 – Espectro eletromagnético com a ampliação do espectro visual da luz.

Ao incidir sobre qualquer objeto, a luz é refletida e captada pelo globo ocular através de um conjunto complexo de elementos que, em seu processo final, são transformados em impulsos nervosos e transmitidos ao cérebro, que os interpreta e trabalha na reconstrução da imagem.

Existem dois tipos de fotorreceptores, também chamados de células sensoriais fotossensíveis, que são responsáveis pela transdução fotoelétrica: cones e bastonetes. Os bastonetes são sensíveis, de uma forma ponderada, a todos os comprimentos de onda do espectro visível, com uma melhor resposta à intensidade da luz. Estão adaptados para a visão em ambientes com baixa iluminação (visão noturna ou escotópica), mas não são muito capazes de distinguir a cor.

Por outro lado, os cones são sensíveis à altos níveis de iluminação e permitem a distinção das cores e detalhes, com melhor resposta na visão diurna. Os cones são divididos em três tipos: vermelho, verdes e azuis.

Todos os estímulos captados pelos bastonetes e cones são enviados a um neurônio, responsável por transportar esse estímulo para diferentes áreas do cérebro, como córtex visual e hipotálamo. O cérebro, por sua vez, recebe milhares dessas terminações nervosas, junta todas as informações e trabalha na reconstrução da imagem.

Como o processamento de cor no olho humano é baseado na amostragem das faixas de vermelho, verde e azul (teoria dos três estímulos ou tricromática – Young-Helmholtz), é natural que se busque uma representação tridimensional do espaço de cor cuja base de cores primárias seja constituída por três cores nesta mesma faixa.

2.2. Sistema de Cores

A representação das cores pode ser dividida em dois principais processos: aditivo e subtrativo. No processo aditivo, duas ou mais ondas de luz são combinadas, formando um raio de luz resultante. Neste processo, a mistura acontece na emissão da luz, onde é possível se gerar todas as cores do espectro com apenas três cores primárias: vermelho, verde e azul (em inglês *Red, Green e Blue - RGB*). A síntese das três cores resulta na cor branca.

Já o processo subtrutivo baseia-se no uso de filtros ou corantes que tem por objetivo filtrar determinados comprimentos de onda, resultando numa nova cor. Este conceito é utilizado em projetores ou na pigmentação de impressoras, por exemplo. Neste processo, a mistura ou subtração acontece com outras três cores, também chamadas de primárias para esse sistema: ciano, magenta e amarelo (em inglês *Cyan, Magenta and Yellow – CMY*). A síntese dessas três cores resulta na cor preta.

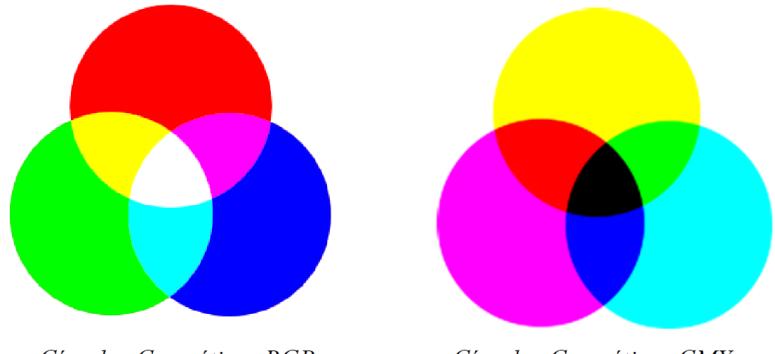


Figura 2.2 – Círculos cromáticos dos sistemas de cores RGB e CMY, respectivamente.

2.3. Modelo RGB

Um modelo de cor é uma representação tridimensional na qual cada cor é posicionada em um determinado ponto do sistema de coordenadas.

O modelo RGB é um sistema aditivo que baseia na captação de cor pelo sistema visual humano (teoria dos três estímulos). É representado pela combinação de três cores primárias: vermelho, verde e azul. É um sistema que não consegue representar todas as cores do espectro visível e que também não é linear com a percepção visual. No entanto, é amplamente utilizado na aquisição, manipulação e reprodução de imagens digitais.

2.4. Imagem Digital

Quando uma imagem é capturada por uma câmera, a luz atravessa as lentes e é projetada em um sensor, que consiste num transdutor fotossensível, capaz de converter energia luminosa sobre um determinado ponto da imagem, em uma carga elétrica.

Quanto mais intensa for essa luminosidade, maior é a corrente elétrica, que é convertida em tensão e lida e transformada em números por meio de um conversor analógico-digital.

Existem duas principais tecnologias que podem ser utilizadas para este sensor em uma câmera: CCD (*Charge-couple device*) e CMOS (*Complementary Metal-oxide Semiconductor*). Em ambas, o sensor puramente não é capaz de reconhecer a cor. Então, faz-se necessário a utilização de filtros na frente dos sensores para que seja possível captar a tonalidade da cor de cada ponto. Estes filtros podem estar projetados para captação de cor pelos sistemas RGB ou CMY, e

são usualmente chamados de “filtro de *Bayer*”. Consiste em uma grade de pontos vermelhos, verdes e azuis, dispostos nas sequências R+G e G+G. Uma vez que o olho humano é mais sensível ao verde que às duas outras cores, a proporção de cores vermelho, verde e azul é de 1:2:1, respectivamente. Após atravessar o filtro, as cores são “separadas” e captadas em cada um dos pixels do sensor.

Para a composição da imagem colorida é necessário um processo de interpolação das três matrizes de cores, que determina e quantifica um valor final para cada pixel utilizando funções matemáticas.

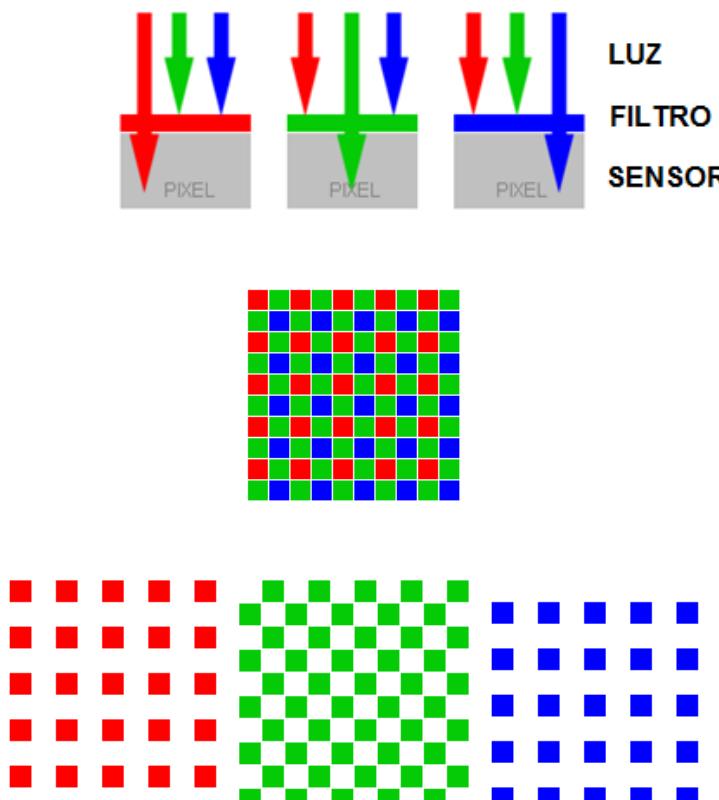


Figura 2.3 – Processo de separação das cores através do “filtro de *Bayer*”; a luz é filtrada e cada uma das três cores (RGB) incidem sobre sensores específicos para cada cor.

3. DESENVOLVIMENTO

3.1. Câmera do Trem

A frota do Monotrilho da Linha 15-Prata, utilizada como objeto de estudo, é equipada em cada um dos seus trens com um sofisticado sistema de monitoramento de vídeo integrado, dotado de 32 câmeras e 2 equipamentos trabalhando em esquema redundante, denominados VCCU – Unidade de Controle de Comunicação. Ambos possuem um sistema VDVR – Registrador de Vídeo Digital do Veículo – integrado, conforme apresentado na Figura 3.1.

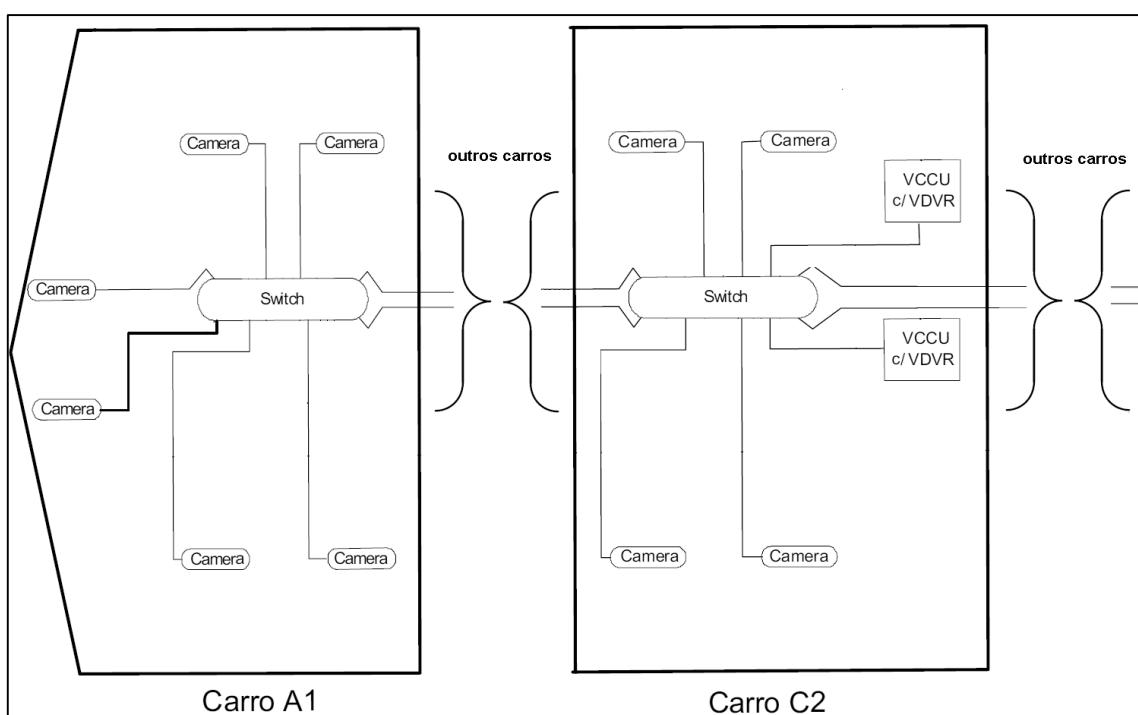


Figura 3.1 – Diagrama de interligação do sistema de monitoramento do trem. Em destaque a configuração de um carro de extremidade (A1) e o carro intermediário (C2), onde estão localizadas as VCCU's. Fonte: adaptado da documentação do Projeto Monotrilho Linha 15-Prata.

A função principal da VCCU é realizar a gestão e estabelecer a comunicação necessária com todas as câmeras do trem por meio da rede de comunicação *Ethernet*. Durante todas as situações de operação do trem as imagens obtidas das câmeras, por intermédio da VCCU, são registradas no VDVR em formato de mídia digital AVI – *Audio Video Interleave*.

As câmeras do trem não desempenham as mesmas funções; existem conjuntos específicos de câmeras configuradas de acordo com sua proposta de utilização. As câmeras voltadas para o interior são aquelas configuradas para monitorar a movimentação interna de cada carro do trem, incluindo suas portas; já as câmeras voltadas para o engate são aquelas configuradas para serem utilizadas para monitorar a região de engate do trem em situações de acoplamento; e, por fim, as câmeras voltadas para o exterior são aquelas configuradas para registrar as atividades externas na via. Para o estudo proposto, as câmeras mais adequadas para obtenção das imagens da via são àquelas voltadas para o ambiente externo.

Fazem parte do conjunto de câmeras voltadas para o exterior apenas duas câmeras do trem e estão posicionadas no teto dos carros de extremidade, com sua lente apontada para a direção do vidro frontal, conforme ilustra a Figura 3.2. As câmeras externas são de fabricante e modelo equivalentes e apresentam as características técnicas indicadas na Tabela 3.1.

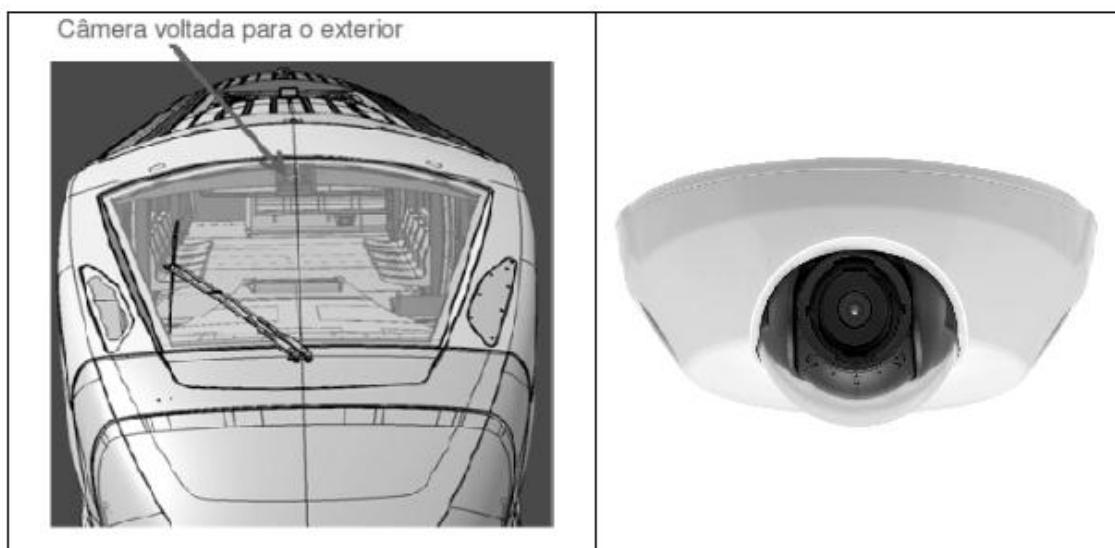


Figura 3.2 – Ilustração indicando a localização da câmera voltada para o exterior do trem e à direita o seu detalhe. Nesta configuração as imagens capturadas podem ser utilizadas para monitoramento das condições à frente da via, local onde existe a movimentação do trem.

Para obtenção das imagens das câmeras, pode-se extrair por meio de uma porta USB – *Universal Serial Bus* – disponível na VCCU, ou por meio de um notebook de manutenção com um *software* dedicado instalado e conectado à rede do trem, onde por protocolo FTP – *File Transfer Protocol* – os vídeos ficam disponíveis para acesso e *download* diretamente do servidor da VCCU.

Por questões relacionadas à capacidade de armazenamento do VDVR, as câmeras externas do trem estão configuradas pela VCCU para trabalharem com algumas de suas características abaixo dos valores nominais apresentados na Tabela 3.1. As câmeras externas estão com valor de taxa de quadros ajustados para operar a 12 quadros por segundo (ou *frames per second – fps*). Entretanto, para fins de avaliação de desempenho dos algoritmos, alguns vídeos foram configurados e extraídos com taxas maiores, como 30 quadros por segundo, e resoluções mais baixas como 700 x 500 pixels.

Ao longo do período de aquisição de dados em campo foi possível extrair 14 vídeos com média de tempo de execução de 1 hora e 25 minutos, representando um total de mais de 6 GB de dados capturados. A Figura 3.3 apresenta uma imagem de um quadro de vídeo capturado pela câmera do trem.



Figura 3.3 – Imagem extraída de um quadro de vídeo capturado pela câmera externa do trem.

Esta imagem foi extraída de um vídeo com resolução de 800 x 600 com taxa de 12 fps - *frames per second*.

Tabela 3.1 – Características técnicas das câmeras externas do trem.

Parâmetro	Descrição
Modelo	AXIS M3113-R
Sensor de imagem	1/4" progressive scan RGB CMOS
Lentes	2,8 mm, 66º visão horizontal, F2.0
Sensibilidade de luz	1 – 100000 lux, F2.0
Tempo de exposição	1/24500 s até 1/6 s
Ângulo de ajuste da câmera	Horizontal ±30º, Vertical 0-90º, Rotação ±180º
Compressão de vídeo	H.264 (MPEG-4 Part 10/AVC) Motion JPEG
Resolução	800x600 até 160x90
Taxa de quadros	Até 30 fps (em todas as resoluções)
Rede	IPv4/v6, HTTP, HTTPSc, SSL/TLS, QoS Layer 3 DiffServ, FTP, CIFS/SMB, SMTP, Bonjour, UPnP™, SNMPv1/v2c/v3 (MIB-II), DNS, DynDNS, NTP, RTSP, RTP, TCP, UDP, IGMP, RTCP, ICMP, DHCP, ARP, SOCKS
Índice de proteção	IP67
Memória	128 MB RAM, 128 MB Flash
Alimentação	Através de Ethernet IEEE 802.3af/802.3at Tipo 1, Classe 1, máx 3,2W
Condições de operação	- 25ºC a 50ºC Umidade 15 – 100% RH (condensação)
Peso	250 g

Para se estudar o comportamento da imagem da câmera em diversas situações, estão apresentadas da Figura 3.4 até a Figura 3.15 algumas situações identificadas como críticas. Dentre elas, estão mostradas as interferências externas, como: sombras provocadas pelas estruturas laterais ou edificações próximas; ofuscamento por falta de luminosidade ou chuva. Mesmo não coletadas as imagens, sabe-se que temporais ou nebulosidades poderão ocorrer e com certeza comprometerão a visibilidade das imagens.



Figura 3.4 – Imagem da via normal, sem nenhuma interferência externa.



Figura 3.5 – Imagem da via com a presença de sombra provocada pelas edificações dos arredores.



Figura 3.6 – Imagem da via com a presença de sombra provocada pela cobertura da estação Vila Prudente (Linha 15-Prata)



Figura 3.7 – Imagem da via com curva suave à esquerda.



Figura 3.8 – Imagem da viga-guia metálica do AMV (notar a diferença de cor e sombra de estruturas vizinhas).



Figura 3.9 – Imagem da via com a presença de sombra provocada pela cobertura da estação Oratório (Linha 15-Prata)



Figura 3.10 – Imagem da via com a presença de sombra provocada por um trem em direção oposta, na via ao lado.



Figura 3.11 – Vista do trem situado na oficina de manutenção, iluminada, com visão para área externa durante a noite.



Figura 3.12 – Imagem da via dentro do patio de manutenção durante a noite e sem iluminação.



Figura 3.13 – Imagem da via dentro do patio de manutenção durante a noite, sem iluminação com curva suave à direita.



Figura 3.14 – Imagem da via durante a noite e sem iluminação, com a lente molhada pela chuva.



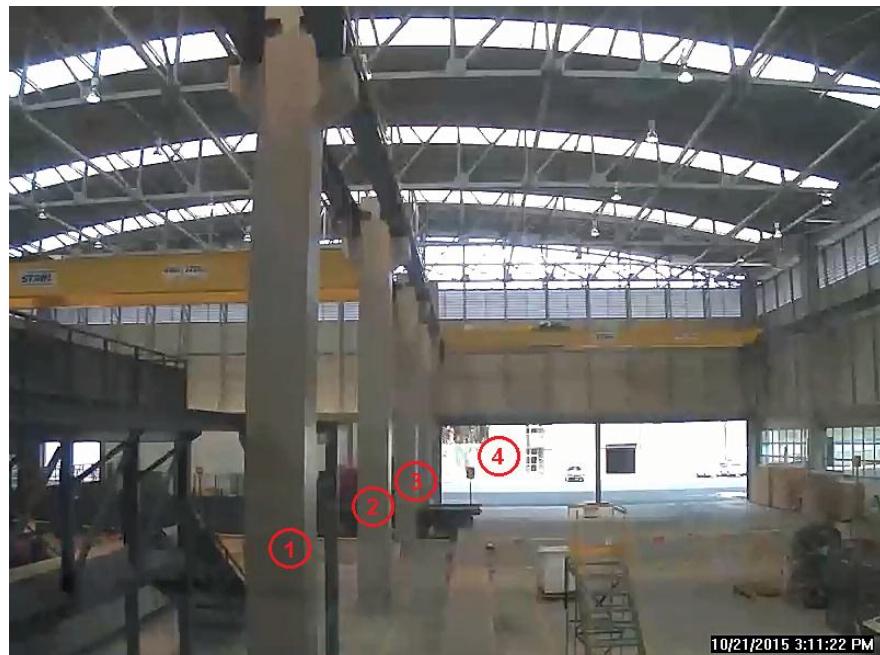
Figura 3.15 – Imagem da via durante a noite, sem iluminação, com a presença das luzes de sinalização do AMV.

3.2. Relação entre a distância real e a distância em pixels

Devido à natureza monocular intrínseca ao sistema de formação de imagens de uma câmera digital (Ballard e Brown, 1982), ao observar as imagens da câmera, é possível verificar que as informações estão projetadas sob um plano em perspectiva, cujo ponto de fuga localiza-se no centro da imagem, facilmente visível na Figura 3.3. Isto implica que, para relacionar a medida de distância à frente do trem na viga, região central inferior na tonalidade cinza clara, é necessário estabelecer uma relação entre a posição dos pixels verticais e a informação de profundidade da imagem, representada pela projeção da viga em perspectiva, e para a realização desta tarefa, fez-se necessário propor o método apresentado a seguir.

Para estabelecer a relação da posição vertical dos pixels com a distância em profundidade da imagem, foi capturado um vídeo onde o movimento do trem foi interrompido em dois pontos equidistantes e conhecidos da via. Este processo foi realizado movimentando-se o trem em direção ao final da linha de estacionamento do Pátio Oratório.

A distância física entre os dois pontos de referência adotados, pilastras de sustentação da cobertura do bloco de manutenção, foi cuidadosamente medida com o auxílio de uma trena e, a partir daí, iniciou-se a movimentação do trem em sentido ao primeiro ponto de referência. Quando a extremidade da câmera externa do trem alcançou o primeiro ponto de referência, o seu movimento foi interrompido por alguns instantes para de criar um marco de parada no movimento do trem para permitir uma localização mais fácil do quadro a ser capturado no vídeo final. E, em seguida, o mesmo procedimento foi repetido para o segundo ponto de referência. Os dois quadros, representados na Figura 3.16 (a) e (b), foram então extraídos do vídeo final e analisados para a determinação da relação pixels-distância.



(a)



(b)

Figura 3.16 – Imagens extraídas do vídeo com o trem parado a (a) 12,0 metros e a (b) 2,8 metros da placa de fim de via, marcação 4 da imagem.

Captura realizada de vídeo com resolução de 800 x 600 pixels.

As marcações 1, 2 e 3, indicadas na Figura 3.16 (a) e (b), têm o objetivo de facilitar a determinação do deslocamento do trem sobre a via. Conforme medidas de campo a distância entre pilastras é 9,20 metros e o poste de sustentação da placa de fim via possui uma distância de 2,8 metros à frente da pilastra com a indicação de

marcação 3. Com as imagens extraídas, o próximo passo foi verificar por meio de um software utilitário gráfico, a contagem de pixels verticais da base da imagem até o ponto do poste da placa de fim de via. Este procedimento está representado na Figura 3.17, onde a linha verde representa a quantidade de pixels verticais da base da imagem até a base do poste.



Figura 3.17 – A linha em verde representa a quantidade de pixels verticais da base da imagem até a base do poste da placa de fim de via. Os quadros à esquerda e à direita são respectivamente as ampliações das imagens capturadas a 12,0 metros e 2,8 metros da placa.

Ao final do procedimento foram contabilizados 129 pixels para a imagem capturada a 12,0 metros da placa e 39 pixels para a imagem capturada a 2,8 metros da placa. Para aferição dos resultados obtidos foi feita uma verificação tentando estimar quantos pixels verticais, calculados por meio da variável x_1 , são necessários para representar 12,0 metros utilizando a relação 39 pixels por 2,8 metros:

$$\frac{x_1}{12} = \frac{39}{2,8} \Rightarrow x_1 \approx 168 \text{ pixels}$$

Após este resultado, uma segunda verificação foi feita, mas em sentido inverso, tentando estimar quantos pixels verticais, calculados por meio da variável x_2 , são necessários para representar 2,8 metros utilizando a relação 129 pixels por 12,0 metros:

$$\frac{x_2}{2,8} = \frac{129}{12} \Rightarrow x_2 \approx 30 \text{ pixels}$$

Os erros das estimativas em relação aos valores obtidos na contagem de pixels das imagens da Figura 3.16, x_1 e x_2 , foram então determinados:

$$\varepsilon_{x_1} = \frac{(39 - 30)}{30} = 30\%$$

$$\varepsilon_{x_2} = \frac{(168 - 129)}{168} = 30\%$$

Ao analisar os erros obtidos foi concluído que os resultados não estavam satisfatórios e era necessário reavaliar o método empregado para encontrar uma relação melhor. Após nova análise mais aprofundada, foi identificado que existia uma outra consideração, a altura da câmera em relação à horizontal do plano da via, que não foi considerada nas primeiras aproximações. A Figura 3.18 apresenta a representação do modelo melhorado, considerando a altura da câmera nas tomadas de distâncias de 12,0 metros e 2,8 metros.

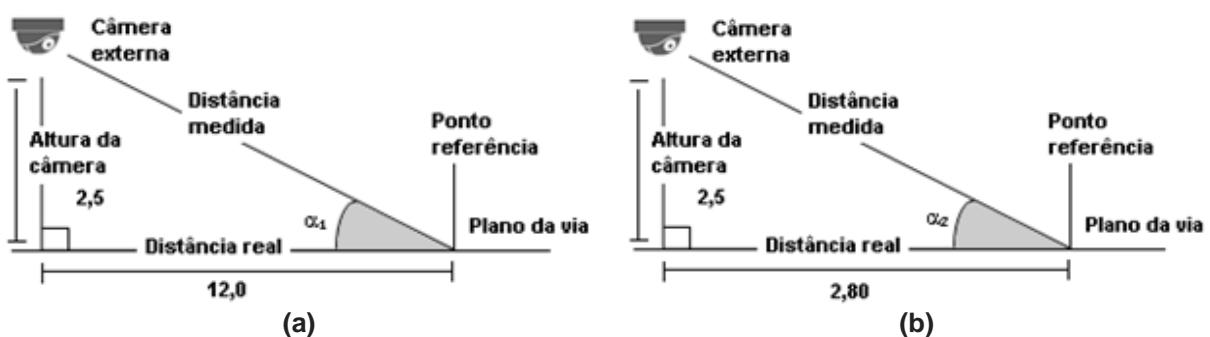


Figura 3.18 – Modelo incorporando a altura da câmera na determinação da relação pixel por metro do sistema. Na representação são considerados os ângulos de projeção em relação ao plano da via nos cenários onde a distância ao ponto de referência, ou placa de fim de via, é (a) 12,0 metros e (b) 2,80 metros.

A partir do modelo, incorporando a altura da câmera, é possível obter os ângulos α_1 e α_2 :

$$\alpha_1 = \tan^{-1}\left(\frac{2,5}{12,0}\right) = 11,8^\circ$$

$$\alpha_2 = \tan^{-1}\left(\frac{2,5}{2,8}\right) = 41,8^\circ$$

Aplicando os fatores de correção, cosseno dos ângulos de projeção, aos valores obtidos na contagem de pixels das imagens da Figura 3.16, obtemos os valores de referência ajustados:

$$y_{ref1} = \cos(\alpha_1) \cdot 129 = 127 \text{ pixels}$$

$$y_{ref2} = \cos(\alpha_2) \cdot 39 = 29 \text{ pixels}$$

Para aferição dos novos valores de referência foi feita uma verificação tentando estimar quantos pixels verticais são necessários para representar 12,0 metros utilizando a relação corrigida, y_{ref2} , de 29 pixels por 2,8 metros:

$$\frac{x'_2}{12} = \frac{29}{2,8} \Rightarrow x'_2 \approx 124 \text{ pixels}$$

O mesmo procedimento foi aplicado para estimar quantos pixels verticais são necessários para representar 2,8 metros utilizando a relação corrigida y_{ref1} :

$$\frac{x'_1}{2,8} = \frac{127}{12} \Rightarrow x'_1 \approx 30 \text{ pixels}$$

Os novos erros das estimativas foram recalculados:

$$\begin{aligned}\varepsilon'_{x_1} &= \frac{(30 - 29)}{29} = 3,3\% \\ \varepsilon'_{x_2} &= \frac{(127 - 124)}{127} = 2,4\%\end{aligned}$$

Avaliando os valores dos novos erros, ε'_{x1} e ε'_{x2} , é possível concluir que a incorporação da altura no modelo de cálculo apresentou uma significativa melhora nos resultados apresentados anteriormente. O passo final é escrever a equação que permite calcular a quantidade de pixels verticais a serem considerados na imagem em função da distância necessária para observação da obstrução da via, assumindo-se que, quando a distância é nula à frente do trem, o número de pixels verticais a serem considerados também deve ser nulo:

$$y = \left(\frac{y_{ref1} - y_{ref2}}{12,0 - 2,80} \right) \cdot d \Rightarrow y = 10,6 \cdot d$$

onde y representa a quantidade de pixels verticais e d representa a distância em metros a ser observada na imagem.

4. CONCLUSÃO

Este trabalho propôs uma abordagem mais detalhada do sistema de visão e captura de imagem da câmera embarcada do trem, parte integrante do sistema de reconhecimento das condições de desalinhamento de via baseado na análise das imagens da câmera (Sistema de Detecção de Desalinhamento de Via – SDDV).

No decorrer do desenvolvimento, foram avaliadas as características das imagens capturadas em distintos cenários de exposição. Durante o dia, notou-se que, quase na maioria dos trechos, há presença de sombras provocadas por: objetos, postes ou estruturas metálicas ao lado e acima da via; trens circulando na via adjacente; edificações próximas ao traçado do monotrilho; coberturas das plataformas das estações. Essas sombras possivelmente prejudicam o tratamento da imagem que antecedem os processos de detecção de borda da viga, objeto de estudo de outra monografia. Essa característica não pode ser eliminada da imagem original, uma vez que se trata da propriedade da via do monotrilho, e exige que haja processos de tratamento da imagem capturada.

Durante o período da noite, e regiões sem a presença de nenhuma outra fonte de iluminação, notou-se que, embora a imagem a olho nu apresente menos visibilidade, há menos interferência externa dos arredores sobre a superfície da via. Desta forma, processos de tratamento de imagem podem obter melhores resultados na identificação das bordas da viga, pois as características de gradiente de cor são mais evidentes. No entanto, durante a noite foi também verificado a presença de água na lente da câmera, provinda de chuva, o que comprometeu a qualidade da imagem. É evidente que este fato independe se imagem foi coletada de dia ou de noite, pois quaisquer condições de intempéries, seja chuva ou neblina, irão comprometer a visibilidade, a claridade e a nitidez da imagem capturada.

Verificou-se que a máxima taxa de quadros que a câmera embarcada do trem dispõe – 30 fps – atenderia melhor ao processo de busca por trechos descontínuos da via. Mas, devido ao reduzido espaço de armazenamento virtual do servidor do VCCU, as câmeras foram previamente configuradas para uma taxa reduzida – 12 fps. Como sugestão, para aumentar a efetividade do sistema, a solução seria reconfigurar as câmeras, aproveitando-se ao máximo ao que o trem dispõe.

Evidentemente, deve-se preocupar em aumentar a capacidade de armazenamento do servidor da VCCU.

Com o intuito de determinar uma relação entre a distância real da via e a distância em pixels da imagem, chegou-se em uma relação de aproximação que apresentou erro máximo de 3,3%. Essa relação é necessária no desenvolvimento dos processos de tratamento de imagem e detecção de bordas, que dependem das distâncias máximas para frenagem de emergência e determinação da região de interesse, necessárias para reduzir o tempo de processamento e garantir tempo de resposta suficiente para uma frenagem segura. Existem outros métodos e relações que podem melhorar os resultados nos cálculos dos erros da relação de conversão, e que irão contribuir para maior segurança do sistema como um todo.

Há inúmeras vantagens na utilização da câmera embarcada do trem, pois o recurso já está adquirido, disponível e o acesso às imagens capturadas pode ser feito em tempo real de forma bastante simplificada através da rede de comunicação do trem. Avaliar a integridade da via através de um sistema visual pode não substituir, mas sim complementar aos processos de decisão dos sistemas de segurança que hoje monitoram a via e as regiões do AMV. A expectativa é que, gradativamente, os sistemas visuais possam substituir outros sistemas que hoje cumprem seu papel, mas com custos de implantação elevados.

5. REFERÊNCIAS BIBLIOGRÁFICAS

BALLARD, D. H.; BROWN, C. M. **Computer vision**. New Jersey: Prentice-Hall, 1982. 523 p.

HATTORI, T.; KATO, T. Analyses of propagation characteristics by VRML models along railway. In: VEHICULAR TECHNOLOGY CONFERENCE, 60., 2004, Japão. **Proceedings of...**

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **Standard for communications-based train control (CBTC) – IEEE 1474**. New York, 2004. pt.1: Performance and functional requirements.

INSTITUTION OF RAILWAY SIGNAL ENGINEERS – IRSE. **Guideline on glossary of signaling and telecommunications terminology**. London, 2011.

KOUASSI, A. J. **Vehicle technology assessment for baltimore transit connector**. Proceedings of the 2001 IEEE/ASME Joint Railroad Conference, 2001, USA.

LAI, Y-C; BARKAN, C. P. L.; DRAPA, J.; AHUJA, N., HART, J. M.; NARAYANAN, P. J.; JAWAHAR, C. V.; KUMAR, A.; MILHON, L. R.; STEHLY M. Machine vision analysis of the energy efficiency of intermodal freight trains. In: JOINT RAIL AND RAPID TRANSIT – JRRT, 2007, USA. **Proceedings of...** USA: 2006, vol. 221.