

**UNIVERSIDADE DE CAXIAS DO SUL**  
**ÁREA DO CONHECIMENTO DE CIÊNCIAS EXATAS E ENGENHARIAS**

**ERICH DANIEL MORAIS**

**UM ESTUDO SOBRE MODELOS CILÍNDRICOS PARA REGISTRO DE IMAGENS**

**Caxias do Sul**  
**2020**

**ERICH DANIEL MORAIS**

**UM ESTUDO SOBRE MODELOS CILÍNDRICOS PARA REGISTRO DE IMAGENS**

Trabalho de Conclusão de Curso apresentado  
como requisito parcial à obtenção do título de  
Bacharel em Engenharia de Computação na  
Área do Conhecimento de Ciências Exatas e  
Engenharias da Universidade de Caxias do Sul

Orientador:  
Prof. Dr. Guilherme Holsbach Costa

**Caxias do Sul**  
**2020**

DANIEL MORAIS, ERICH

Um estudo sobre modelos cilíndricos para registro de imagens / ERICH DANIEL MORAIS; orientador: Guilherme Holsbach Costa, Caxias do Sul: UCS, Centro de Ciências Exatas e Tecnologia, 2020.

40 folhas, 30 cm.

Trabalho de conclusão do curso de Engenharia de Computação — Universidade de Caxias do Sul, Centro de Ciências Exatas e Tecnologia, 2020

Inclui referências bibliográficas.

1. Inspetores de rótulos. Imagem panorâmica. Mosaico de imagens. Registro de imagens. I. Costa, Guilherme H. II. Universidade de Caxias do Sul. III. Um estudo sobre modelos cilíndricos para registro de imagens.

**ERICH DANIEL MORAIS**

**UM ESTUDO SOBRE MODELOS CILÍNDRICOS PARA REGISTRO DE IMAGENS**

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Bacharel em Engenharia de Computação na Área do Conhecimento de Ciências Exatas e Engenharias da Universidade de Caxias do Sul

**Aprovado em 10/12/2020**

**Banca Examinadora**

---

Prof. Dr. Guilherme Holsbach Costa  
Universidade de Caxias do Sul - UCS

---

Eng. Gustavo de Camargo  
Universidade de Caxias do Sul - UCS

---

Prof. Dr. Andre Luis Martinotto  
Universidade de Caxias do Sul - UCS

## AGRADECIMENTOS

Ao meu orientador, Guilherme H. Costa, pela total dedicação, compreensão e comprometimento em me auxiliar não só no desenvolvimento deste trabalho, mas também em meio a vida acadêmica. Obrigado por me ensinar a trilhar o caminho mais recompensador, no qual exige total comprometimento na execução das tarefas realizadas.

Aos meus professores e orientadores de iniciação científica, pela paciência e por nunca terem se recusado a ajudar, mesmo quando não lhes era conveniente.

Aos amigos Anderson, Matheus Nathan, Thiago, Douglas e Mateus Gasparly que trilharam o caminho da graduação ao meu lado, compartilhando conhecimentos e ajudando-me a superar novos desafios constantemente. Um agradecimento especial a Iuri Crestani e Rodrigo Dallagnol, que do primeiro ao último dia da graduação, fizeram a árdua caminhada valer muito mais a pena.

Aos amigos de longa data João Paulo Guedes, Emiliano Piazer e Lucas Matos por sempre estarem ao meu lado, mesmo nos momentos mais difíceis. Um destaque a Aleksandro Gabrielli que teve influência direta na minha escolha profissional.

Aos meus familiares Halisson, Gustavo, Juarez e Deonice que sempre fizeram o possível e o impossível para que eu concluísse o curso, me motivando a seguir em frente perante todas as dificuldades. Espero imensamente poder retribuir tudo que foi feito por vocês durante esse período, sem o apoio e incentivo prestados por vocês, nada disso teria sido possível.

E por fim, a todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

*“To get what you want, you must look beyond what you see.”*

Rafiki, The Lion King

## RESUMO

Na indústria de envase de bebidas é comum a demanda por sistemas de visão computacional chamados de inspetores de rótulos. Esses sistemas verificam o posicionamento dos rótulos afixados em garrafas, permitindo a retirada automática das garrafas que estejam fora dos padrões de qualidade. Entretanto, os inspetores comerciais são importados, possuem alto custo, e, aparentemente, se baseiam em reconhecimento de padrões, requerendo novo treinamento para cada novo modelo de garrafa ou rótulo. Uma possibilidade de solução para verificação do posicionamento dos rótulos é uma abordagem baseada no registro de imagens e na correlação entre imagens. Para tanto, como etapa de pré-processamento, pode-se construir uma imagem panorâmica da superfície das garrafas. Visto isso, neste trabalho é proposto o estudo de uma solução para geração de imagens panorâmicas da superfície de objetos cilíndricos.

**Palavras-chave:** Inspetores de rótulos. Imagem panorâmica. Mosaico de imagens. Registro de imagens.

## **ABSTRACT**

In the beverage filling industry, demand for computer vision systems called label inspectors is common. These systems check the positioning of the labels affixed to the bottles, allowing the automatic removal of bottles that are outside the quality standards. However, commercial inspectors are imported, expensive and seemingly based on pattern recognition, requiring new training for each new bottle or label model. A possible solution to check the positioning of the labels is an approach based on the registration of images and the correlation between images. For this, as a pre-processing step, a panoramic image of the bottles' surface can be constructed. In view of this, this work proposes the study of a solution to generate panoramic images of the surface of cylindrical objects.

**Keywords:** Label inspectors. Panoramic image. Image mosaic. Image registration.

## LISTA DE FIGURAS

Figura 1 – Modelo geométrico de imagens panorâmicas. . . . .	11
Figura 2 – Vista superior da disposição das câmeras. . . . .	13
Figura 3 – Diagrama esquemático da distorção causada no eixo horizontal. . . . .	14
Figura 4 – Diagrama da distorção vertical. (a) Imagem sem a distorção; (b) Imagem distorcida. . . . .	15
Figura 5 – Forma geométrica de uma elipse representada no plano cartesiano. . . . .	16
Figura 6 – Forma geométrica de uma elipse representando a distorção do eixo vertical. . . . .	16
Figura 7 – Pontos correspondentes em diferentes segmentos de imagem. . . . .	17
Figura 8 – Demonstração de uma transformada geométrica projetiva. . . . .	19
Figura 9 – Destaque dos <i>key points</i> da imagem, obtidos com a utilização do SIFT. . . . .	20
Figura 10 – Exemplo de correspondência entre pontos com a utilização do SIFT. . . . .	21
Figura 11 – Remoção dos <i>outliers</i> com a utilização do RANSAC. . . . .	21
Figura 12 – Diagrama esquemático da plataforma utilizada para a aquisição das imagens. . . . .	23
Figura 13 – Etapas do sistema proposto. . . . .	23
Figura 14 – Pontos que representam as extremidades do rótulo (em amarelo) e a elipse imaginária (em vermelho). . . . .	25
Figura 15 – Aplicação da transformada geométrica projetiva no mapeamento das distorções. . . . .	27
Figura 16 – Esquemático do dimensionamento dos blocos que compõem a imagem. . . . .	27
Figura 17 – Plataforma utilizada para a aquisição das imagens. . . . .	30
Figura 18 – Aplicação do algoritmo para obtenção da distorção. . . . .	31
Figura 19 – Pontos que representam as distorções características desta garrafa. . . . .	32
Figura 20 – Imagem do objeto analisado. . . . .	32
Figura 21 – Resultado da planificação com $N = 5$ . . . . .	33
Figura 22 – Resultado da planificação com $N = 10$ . . . . .	33
Figura 23 – Resultado da planificação com $N = 20$ . . . . .	34
Figura 24 – Resultado da planificação com $N = 30$ . . . . .	34
Figura 25 – Tempo de execução (em segundos) para a quantidade de pontos total atribuída. . . . .	35
Figura 26 – Resultado da panorâmica cortando $45^\circ$ . . . . .	36
Figura 27 – Imagens utilizadas para a formação da panorâmica. . . . .	40

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>10</b>
1.1	OBJETIVOS	12
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>13</b>
2.1	CONSTRUÇÃO DAS IMAGENS	13
2.1.1	Projeções cilíndricas	14
2.1.2	Compensações geométricas	17
2.1.3	Mosaico de imagens	19
<b>3</b>	<b>MÉTODO PROPOSTO</b>	<b>22</b>
3.1	PLATAFORMA DE AQUISIÇÃO	22
3.2	SISTEMA PROPOSTO	23
3.3	PRÉ-PROCESSAMENTO DA IMAGEM	24
3.4	MAPEAMENTO DAS DISTORÇÕES	25
3.5	COMPENSAÇÕES DAS DISTORÇÕES	26
3.6	MONTAGEM DO MOSAICO	28
<b>4</b>	<b>RESULTADOS</b>	<b>30</b>
4.1	PLATAFORMA DE AQUISIÇÃO	30
4.2	IDENTIFICAÇÃO DAS DISTORÇÕES PROJETIVAS	31
4.3	REMOÇÃO DAS DISTORÇÕES	32
4.4	MONTAGEM DA PANORÂMICA	35
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>37</b>
	<b>REFERÊNCIAS</b>	<b>38</b>
	<b>ANEXO A- MONTAGEM DA PANORÂMICA</b>	<b>40</b>

## 1 INTRODUÇÃO

Em meio a grande variedade de opções de produtos vindos do mundo inteiro para o consumidor brasileiro, a indústria nacional que é destinada à produção de bebidas passa por um momento de transformação de conceitos no qual a qualidade deve estar em destaque. Os rótulos têm uma grande influência na decisão de compra do consumidor final. Elementos como cores, formas, materiais, imagens e a linguagem utilizada, podem trazer ao rótulo uma identidade única, influenciando em sentimentos, percepções e valores funcionais aos produtos, que auxiliam e atraem o consumidor na decisão de compra. Em sua maioria, contêm a identidade de uma marca, assim como informações relevantes ao produto, tal como funcionalidades, prazos de validade, avisos e etc.

Quando um rótulo apresenta alguma deformidade na sua impressão ou fixação pode levar ao consumidor uma percepção de descuido ou sentimento pejorativo com relação ao produto, influenciando diretamente na sua decisão de compra. Com isso, as empresas do ramo procuram se desenvolver tecnologicamente com o intuito de minimizar essas deformidades. A solução geralmente adotada é a inspeção dos produtos, visando a eliminação daqueles fora de conformidade. Para isso, as pequenas empresas tendem a utilizar inspeção visual humana, que é lenta e não confiável, visto que não há uma tecnologia automatizada nacional e/ou com custo acessível para esse tipo de aplicação. Já para as empresas de grande porte, tende a ser viável a opção por tecnologias importadas, geralmente com alto valor de mercado.

Uma das alternativas de inspeção automatizada disponíveis no mercado é a tecnologia da empresa alemã EVT (<http://www.evt-web.com/>), que utiliza uma ferramenta CAD (*Computer-Aided Design*) para fazer a análise do rótulo através de imagens adquiridas em tempo-real. As garrafas são submetidas a um processo de medição por imagens e caso não ocorra diferenças consideráveis entre a disposição do rótulo aplicado e um padrão pré-definido, a garrafa é aprovada por estar dentro dos padrões de qualidade da empresa. Um ponto negativo desse método é a necessidade de posicionar as garrafas com o rótulo virado para a câmera no momento da inspeção. Caso isso não ocorra, o sistema pode rejeitar garrafas com padrões de qualidade aceitáveis. Com vistas a evitar um posicionamento preciso das garrafas frente à câmera, alguns sistemas (<https://www.ftsystem.com/english/>) utilizam mais de uma câmera, de forma a adquirir imagens de diversos ângulos da garrafa. Os sistemas comerciais existentes possivelmente buscam, entre essas imagens, o melhor ângulo de observação de cada característica a ser inspecionada.

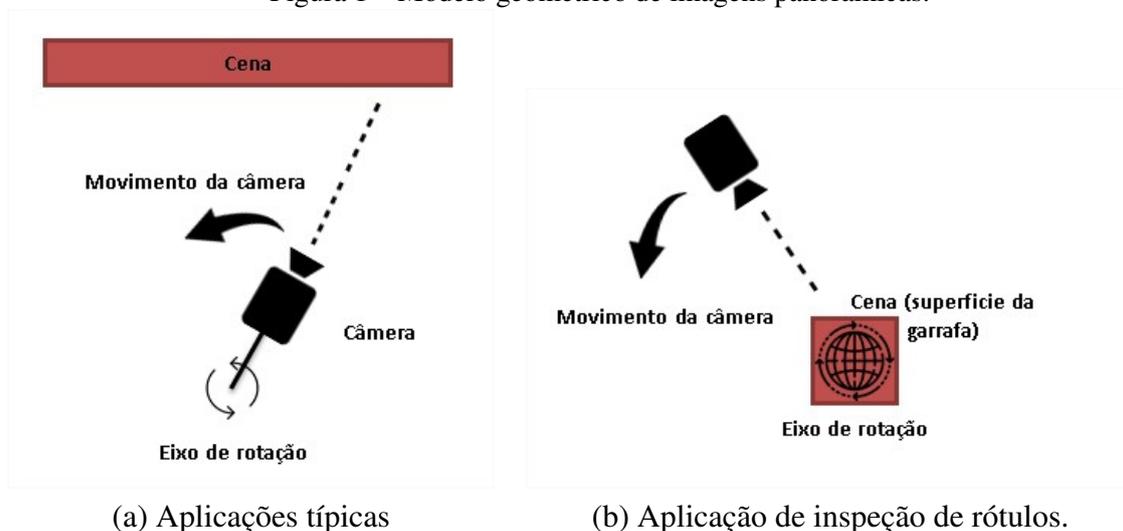
O problema de inspeção de rótulos é pouco abordado na comunidade científica, não gerando resultados consistentes com o assunto quando, por exemplo, as palavras-chave *bottle*, *label* e *inspection*, jargões técnicos usuais na área, são utilizadas nas ferramentas de procura do IEEE Xplore e do Google Scholar. Até onde alcançou a pesquisa bibliográfica que embasa este trabalho, um único artigo foi encontrado com o propósito de inspeção de rótulos (LIN et al., 2013).

Em (LIN et al., 2013) é proposto um sistema de inspeção baseado na utilização de qua-

tro câmeras, que atuam simultaneamente, a fim de capturar imagens em torno do perímetro da garrafa. Essas quatro imagens são utilizadas para compor uma imagem panorâmica que, posteriormente, pode ser utilizada para comparação com uma panorâmica padrão (ainda que o artigo não avance até tal etapa). Para isso, os autores apresentam um modelo trigonométrico para compensação das distorções projetivas presentes em cada uma das cenas. Uma vez feitas essas compensações, as imagens são alinhadas por meio de um algoritmo de *registro*, responsável por estimar o deslocamento relativo entre imagens.

A maneira de se compor uma imagem panorâmica (e, por hipótese, periódica) da superfície da garrafa possibilita que eventuais problemas de posicionamento dos rótulos possam ser detectados por alguma medida de similaridade; por exemplo, estatística, como a correlação. A formação de imagens panorâmicas é um tema amplamente abordado na literatura (LEE; KIM; CHOI, 2001; PARK et al., 2013). Entretanto, o modelo geométrico das aplicações usuais é geralmente o inverso do modelo geométrico da aplicação em questão. As aplicações tradicionais se configuram como uma câmera que é rotacionada em torno de um eixo de fixação, estando ela posicionada ao longo do raio de rotação e direcionada no sentido oposto ao eixo (ver Figura 1 (a)). Na aplicação de inspeção de rótulos, a geometria do problema é similar, mas a câmera fica direcionada no sentido do eixo de rotação (ver Figura 1 (b)).

Figura 1 – Modelo geométrico de imagens panorâmicas.



Fonte: O autor (2020)

Existem outras aplicações com essa mesma geometria, cujos autores propõem soluções semi-supervisionadas (LEE; MAGNENAT-THALMANN, 2000) ou propõem soluções de alto custo de hardware (baseadas em *scanners* 3D) e de processamento (KOVACS et al., 2006), enquanto as linhas de produção de bebidas produzem em velocidades típicas que variam de 2,5 a 15 garrafas por segundo. Além disso, as condições de aquisição em linhas de envase de bebidas são mais facilmente controladas e o produto sob inspeção possui uma geometria substancialmente mais simples que em (LEE; MAGNENAT-THALMANN, 2000; KOVACS et al., 2006), o que potencialmente possibilita o uso de técnicas igualmente mais simples (e portanto

computacionalmente menos custosas) e com resultados mais confiáveis (PARK et al., 2013).

O presente trabalho estuda a aplicação do método proposto em (LIN et al., 2013), porém sem o uso de algoritmos de registro. Para tanto, o deslocamento entre as cenas é assumido conhecido, com base na posição das quatro câmeras.

O presente texto está dividido em cinco capítulos. Este primeiro Capítulo é introdutório e, na próxima seção, ainda são apresentados os objetivos do trabalho. O segundo Capítulo descreve a teoria necessária para realizar o mapeamento das distorções, compensações geométricas e a técnica utilizada por (LIN et al., 2013) para a montagem de um mosaico. O método proposto é apresentado no Capítulo 3. No Capítulo 4 são apresentados e discutidos os resultados obtidos. Por fim, o Capítulo 5 finaliza o trabalho.

## 1.1 OBJETIVOS

### **Objetivo geral**

Gerar uma imagem panorâmica da superfície de um objeto cilíndrico.

### **Objetivos específicos**

1. Investigar uma técnica potencialmente adequada à construção de imagens panorâmicas da superfície de garrafas;
2. Avançar o conhecimento do Núcleo de Inovação e Desenvolvimento em Controle e Automação, da Universidade de Caxias do Sul, a respeito de técnicas de mosaico de imagens;
3. Avançar o conhecimento do Núcleo de Inovação e Desenvolvimento em Controle e Automação, da Universidade de Caxias do Sul, na direção do desenvolvimento de um inspetor de rótulos para linhas de envase de bebidas.

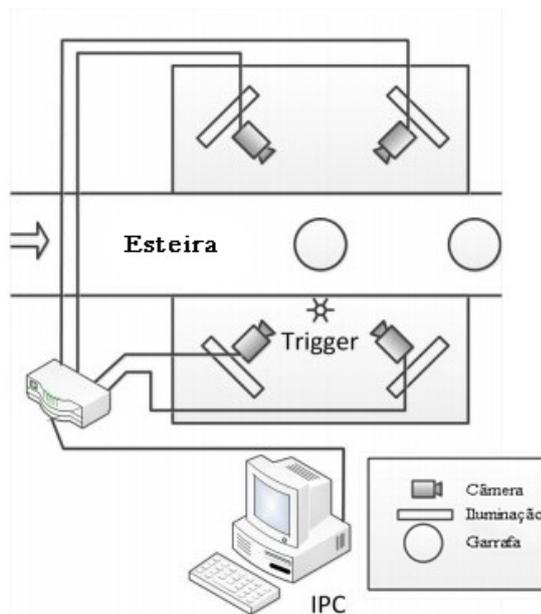
## 2 FUNDAMENTAÇÃO TEÓRICA

Para uma série de aplicações, a maneira mais adequada de se obter imagens da superfície de um objeto cilíndrico é adquirir fotografias dessa superfície a partir de diferentes ângulos e, posteriormente, realizar a união destas imagens através de algoritmos de registro utilizados em imagens panorâmicas e outros tipos de mosaicos de imagens (PARK et al., 2013). Com a intenção de diminuir deformações causadas pelas distorções projetivas, bem como pelas diferentes perspectivas de projeção geradas pela geometria tridimensional do objeto de interesse, em (LIN et al., 2013) é proposto um algoritmo que corrige essas distorções, utilizando técnicas de trigonometria que proporcionam a identificação de uma coordenada específica da imagem que representa uma distorção por ângulo. Estes pontos são submetidos a uma transformação geométrica de acordo com a sua posição na imagem. Neste capítulo, o método proposto em (LIN et al., 2013), bem como sua aplicação, é detalhado.

### 2.1 CONSTRUÇÃO DAS IMAGENS

A solução proposta em (LIN et al., 2013) se baseia na fusão de imagens obtidas através de ângulos distintos, a fim de ilustrar diferentes regiões da superfície de um cilindro. A Figura 2 representa uma vista superior da disposição entre a garrafa e as quatro câmeras posicionadas com o objetivo de garantir uma visão completa do perímetro do objeto.

Figura 2 – Vista superior da disposição das câmeras.



Fonte: Adaptado de Lin et al. (2013).

Neste sistema, uma esteira realiza o movimento da garrafa que, quando posicionada em frente ao *trigger*, dispara o comando de captura fazendo com que todas as imagens sejam obtidas

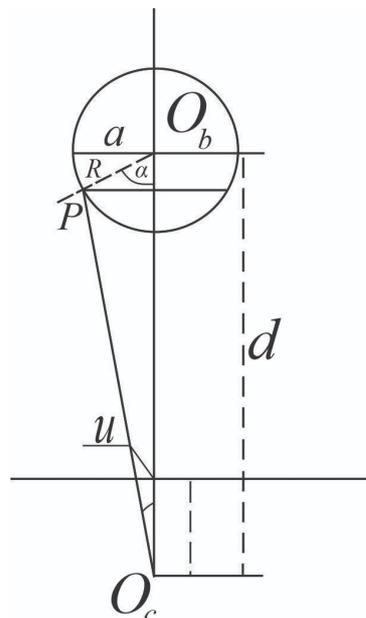
simultaneamente. Por fim, estas imagens são enviadas ao IPC (*Industrial Personal Computer*) que é responsável por processá-las.

Basicamente, três tipos de distorções são encontradas em imagens adquiridas no contexto deste sistema. O primeiro corresponde às distorções projetivas geradas pelo desalinhamento do plano do sensor com o plano da imagem (cena). O segundo diz respeito às distorções projetivas causadas pela projeção de um objeto (cena) tridimensional em um plano (sensor de aquisição). O terceiro diz respeito às distorções geradas pelo conjunto óptico, principalmente as distorções geométricas geradas pela objetiva. O primeiro tipo de distorção é tratado em (LIN et al., 2013) por meio da biblioteca de calibração *Camera Calibration Toolbox* do Matlab® (BOUGUET, 2017). Visando a avaliação de uma solução que dispense processos elaborados de calibração, neste trabalho, a compensação desse tipo de distorção é negligenciada. O mapeamento das demais distorções é abordado na seção seguinte.

### 2.1.1 Projeções cilíndricas

A projeção de um objeto (aproximadamente) cilíndrico no plano do sensor gera distorções nas direções vertical e horizontal. É possível dizer que todas as representações de superfícies curvas em um plano envolvem “contrações” ou “extensões” que podem resultar em distorções ou “rasgos” na tentativa de corrigir estes artefatos através de compensações geométricas (SHEN, 2013). Diferentes técnicas de identificação e representação dessas distorções são utilizadas no sentido de se alcançar resultados que possuam certas propriedades favoráveis para um propósito específico (SHEN, 2013; ZHAO et al., 2013).

Figura 3 – Diagrama esquemático da distorção causada no eixo horizontal.



Fonte: Adaptado de Lin et al. (2013).

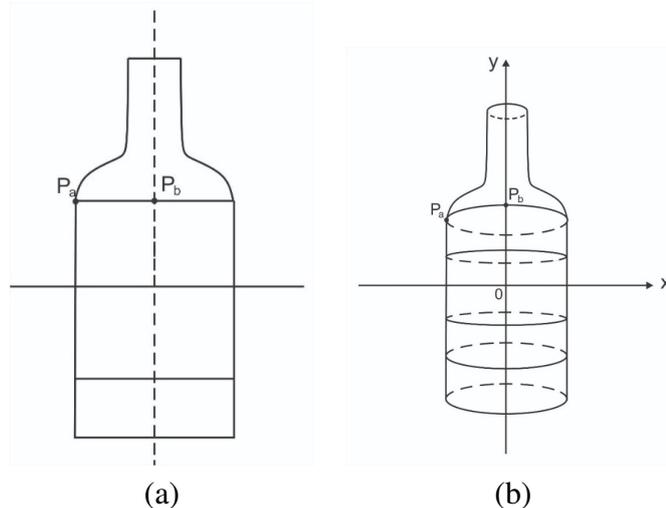
A fim de obter as coordenadas que mapeiam a distorção do eixo  $x$  (horizontal), em (LIN et al., 2013) foi considerado o diagrama esquemático apresentado na Figura 3. Nesse diagrama,  $u$  representa a coordenada  $x$  do ponto  $P$  e  $\alpha$  denota o ângulo entre os segmentos de reta  $\overline{O_bP}$  e  $\overline{O_bO_c}$ . Observando o diagrama é possível estabelecer, através de relações trigonométricas, a função que representa as distorções horizontais, dada por

$$u = a \frac{\sqrt{d^2 - R^2} \sin(\alpha)}{d - R \cos(\alpha)} \quad (2.1)$$

em que  $R$  é o raio do cilindro (garrafa),  $a$  é a distância entre o ponto extremo horizontal (da garrafa) e o seu respectivo centro (na imagem) e  $d$  é a distância entre o centro óptico da câmera,  $O_c$ , e o centro da garrafa,  $O_b$ . Levando em consideração que calibrações intrínsecas à câmera estão sendo negligenciadas (não permitindo o real dimensionamento do ângulo máximo real alcançado pelo sensor), neste trabalho o limite lateral esquerdo da garrafa será considerado ideal, representando  $-90^\circ$  enquanto o seu ângulo de meia volta identifica o limite à direita do rótulo.

Em (LIN et al., 2013) as distorções geradas pela a perspectiva da cena em relação ao eixo vertical são modeladas como uma elipse, assim como é ilustrado na Figura 4, onde a Figura 4 (a) representa o rótulo ideal (sem distorções verticais) e a Figura 4 (b) demonstra o efeito da distorção causado na fotografia. Na imagem em que o rótulo não apresenta distorções, é possível observar que apesar de todos os pontos entre  $P_a$  e  $P_b$  terem a mesma ordenada a distorção projetiva causa uma forma geométrica distinta ao plano da imagem obtida, representada no diagrama ao lado. Na figura distorcida, nota-se que quando o ponto se encontra acima do ponto central da imagem, a elipse tem o segmento de curva para cima e quando o ponto se encontra abaixo, tens o segmento inferior da elipse. Na Figura 4 (b) os pontilhados auxiliam na visualização da forma geométrica completa da elipse.

Figura 4 – Diagrama da distorção vertical. (a) Imagem sem a distorção; (b) Imagem distorcida.

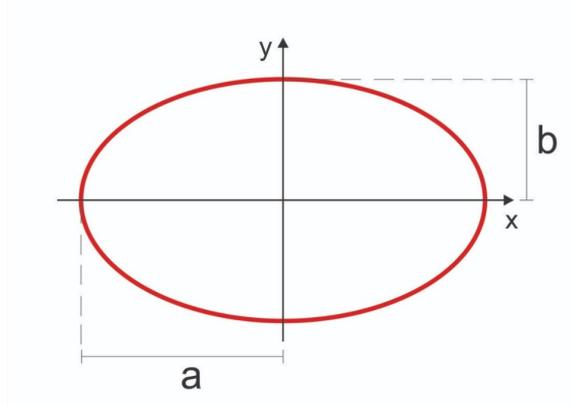


Fonte: O autor (2020).

Para um melhor entendimento da equação que representa as distorções verticais, se vê necessária uma breve abordagem na equação fundamental que representa a elipse no plano cartesiano (ver Figura 5). Em geometria, uma elipse é um tipo de seção cônica, isto é, caso uma superfície cônica seja cortada por um plano que não passe pela sua base e que não ultrapasse as duas folhas do cone, a interseção entre o cone e o plano é uma elipse. As medidas da elipse são dadas pela metade dos eixos maior e menor sendo chamadas, respectivamente, de semi-eixo maior ( $a$ ) e semi-eixo menor ( $b$ ). A equação fundamental da elipse é dada por

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (2.2)$$

Figura 5 – Forma geométrica de uma elipse representada no plano cartesiano.



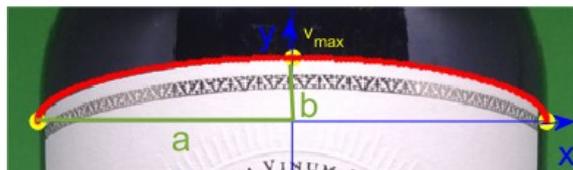
Fonte: O autor (2020).

Com isso, Lin et al. (2013) mapeiam as distorções verticais através da equação descrita a seguir:

$$v = h \frac{R}{d + R} \sqrt{1 - \frac{(d^2 - R^2) \sin^2(\alpha)}{[d - R \cos(\alpha)]^2}} \quad (2.3)$$

em que  $h$  representa a distância entre o centro do rótulo e sua ordenada máxima (na imagem), chamada de  $v_{max}$ . Uma representação da elipse que identifica uma distorção vertical pode ser vista na Figura 6.

Figura 6 – Forma geométrica de uma elipse representando a distorção do eixo vertical.



Fonte: O autor (2020).

Com as coordenadas  $u$  e  $v$  que representam as distorções horizontais e verticais respectivamente identificadas, é possível expandir (realizar as compensações geométricas) a imagem a

partir dos pontos mapeados pelo ângulo  $\alpha$ . Realizando este procedimento com cada uma das imagens, serão formadas no total quatro novas imagens planificadas que, ao se unirem, representarão o perímetro total da garrafa.

### 2.1.2 Compensações geométricas

Em (LIN et al., 2013) não é detalhado o método adotado para realizar as compensações geométricas. Portanto, nesta revisão são consideradas transformações geométricas a partir do mapeamento das distorções identificadas na imagem. Segundo (GONZALEZ; WOODS, 2000), uma transformação geométrica consiste na operação de uma transformação espacial na imagem, sendo essa, responsável pela "reorganização" dos pixels através de regras pré-estabelecidas. Para deixar o procedimento mais claro, será mencionado o exemplo utilizado em (GONZALEZ; WOODS, 2000):

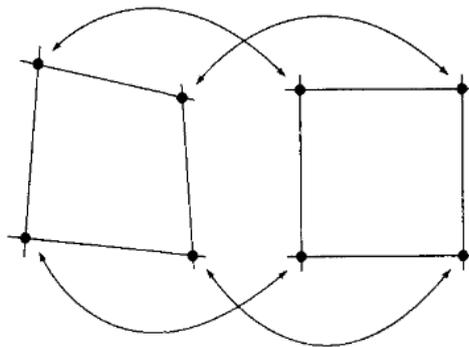
Supondo que uma imagem denominada  $f$  com coordenadas de pixels definidas por  $(x, y)$  seja submetida a uma transformação geométrica para se tornar uma nova imagem  $g$ , onde suas novas coordenadas serão dadas por  $(x', y')$ , a transformação é dada pela seguinte notação:

$$\begin{aligned} x' &= r(x, y) \\ y' &= s(x, y) \end{aligned} \quad (2.4)$$

em que  $r(x, y)$  e  $s(x, y)$  são as transformações que produzem uma nova imagem  $g(x', y')$ . Por exemplo, se  $r(x, y) = x/2$  e  $s(x, y) = y/2$ , a distorção pode ser descrita como um "achatamento" de  $f(x, y)$ , reduzindo seu tamanho pela metade em ambos os eixos.

O método mais utilizado para modelar essa transformação é a seleção de *tiepoints*, que são coordenadas conhecidas na imagem distorcida e na imagem original (GONZALEZ; WOODS, 2000). Na Figura 7 é possível visualizar um paralelogramo (figura distorcida) e um quadrado (figura corrigida), onde os *tiepoints* são destacados pelos vértices.

Figura 7 – Pontos correspondentes em diferentes segmentos de imagem.



Uma (matriz de) transformação comumente utilizada e que é consistente com o problema abordado neste trabalho é a (matriz de) transformação projetiva. Essa transformação, causa o efeito de inclinação sob o plano da imagem, onde normalmente as linhas paralelas convergem para um "ponto de fuga" (ver Figura 8). A equação que descreve uma transformação geométrica na imagem é dada por (WOLBERG, 1990) como

$$[x', y', w'] = [u, v, w]\mathbf{T} \quad (2.5)$$

em que

$$\mathbf{T} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

e  $[x', y', w']$  é a coordenada homogênea<sup>1</sup> de  $\mathbf{p} = [x, y]$ .

Diferente de uma transformação afim (mudança de escala, movimento de rotação ou translação), os valores da última coluna são diferentes de zero (WOLBERG, 1990). Portanto,  $a_{13}$  e  $a_{23}$  influenciam diretamente no "ponto de fuga", de maneira que, quanto maiores são os seus valores, mais o ponto de fuga se aproxima da origem e, portanto, as linhas paralelas parecem convergir mais rapidamente, proporcionando uma inclinação maior no plano da imagem, causando uma "sensação" de profundidade na cena (ver Figura 8). Portanto, o mapeamento (fora da forma homogênea) entre a imagem transformada e a imagem original é dado por

$$\begin{aligned} x &= \frac{x'}{w'} = \frac{a_{11}u + a_{21}v + a_{31}}{a_{13}u + a_{23}v + a_{33}} \\ y &= \frac{y'}{w'} = \frac{a_{12}u + a_{22}v + a_{32}}{a_{13}u + a_{23}v + a_{33}} \end{aligned} \quad (2.6)$$

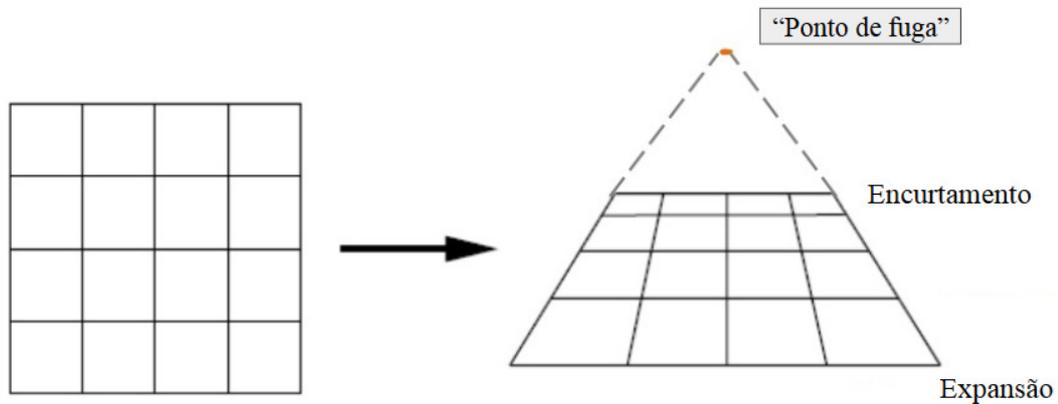
Observando que a matriz  $\mathbf{T}$  possui dimensões  $3 \times 3$ , é possível afirmar que sua normalização resulta em  $a_{33} = 1$ , garantindo oito "termos livres" na matriz. Com isso, é possível estabelecer uma transformação geométrica projetiva a partir dos quatro pontos conhecidos (*tiepoints*) na imagem original e quatro pontos na imagem distorcida.

Pode-se ressaltar a afirmação de (GONZALEZ; WOODS, 2000) novamente, em que a utilização de *tiepoints* para modelar a transformação entre ambas as imagens pode ser estabelecida. Em (WOLBERG, 1990) e (HECKBERT, 1989) os autores exploram a fundo métodos que aceleram os procedimentos de transformada geométrica, considerando comportamentos conhecidos entre as coordenadas das imagem distorcida e da imagem retificada e melhoria de desempenho, entre outros fatores.

---

<sup>1</sup>Coordenadas homogêneas são um sistema de coordenadas utilizadas em geometria projetiva, assim como coordenadas cartesianas são utilizadas em geometria Euclidiana. A vantagem delas é que pontos, incluindo pontos no infinito, podem ser representados por coordenadas finitas. Formulações que envolvem coordenadas homogêneas são geralmente mais simples e simétricas que as respectivas formulações para coordenadas cartesianas. (CONNOR; ROBERTSON, 1997)

Figura 8 – Demonstração de uma transformada geométrica projetiva.



Fonte: O autor (2020).

### 2.1.3 Mosaico de imagens

A construção de uma imagem panorâmica é uma técnica que reúne múltiplas imagens adquiridas a partir de diferentes vistas. O principal objetivo é formar uma imagem única e abrangente, também chamada de mosaico (SHUM; SZELISKI, 2000). Para que o mosaico tenha uma qualidade considerável, a união das imagens (também conhecida por “fusão”) é um dos passos mais importantes. Esse tipo de técnica se baseia em premissas matemáticas que demandam, por exemplo, a existência de semelhanças e sobreposições consideráveis entre regiões das diferentes imagens a serem fundidas. Entretanto, é comum que ocorram variações geométricas nas áreas de sobreposição das imagens ou violações nessas premissas durante o processo de aquisição, acarretando em artefatos degradantes, como regiões desfocadas ou costuras visíveis entre as imagens (GRACIAS et al., 2009).

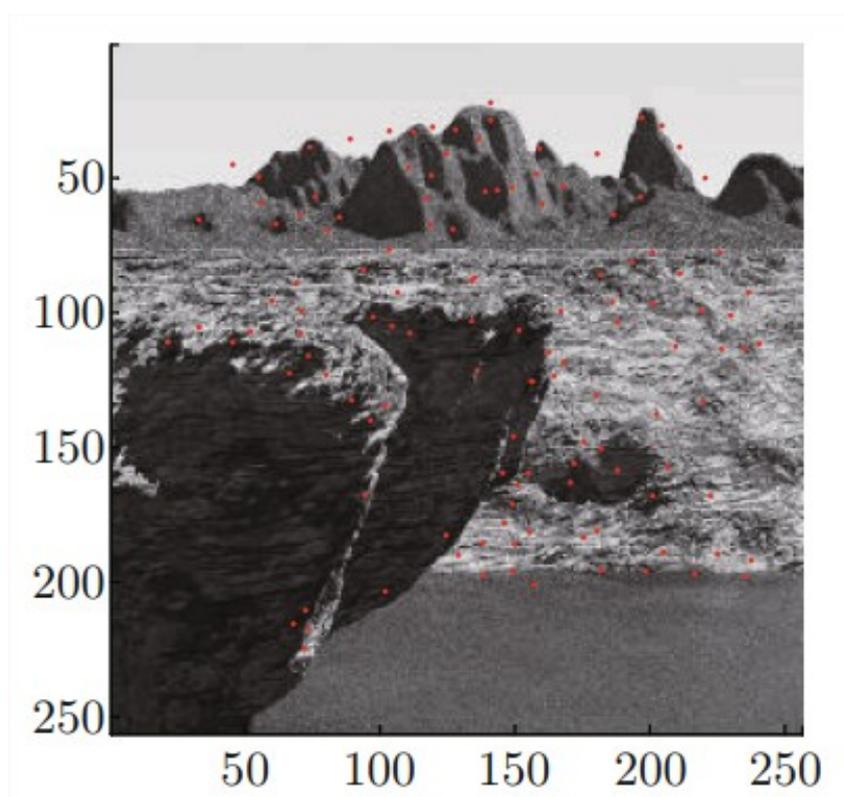
Em (LEVIN et al., 2004), as técnicas de mosaico são divididas em duas abordagens: suavizações entre transições; e localizações ideais de costura. O trabalho proposto em (LIN et al., 2013) segue a segunda abordagem, visando obter um resultado mais preciso na união das imagens. Dessa forma, foi empregada uma técnica de registro de imagens que permite contornar os problemas de mudança de escala, variações de iluminação e a possibilidade da garrafa ser apresentada para as câmeras rotacionada (em diferentes posições angulares). Dentre as diversas técnicas disponíveis para a detecção de características comuns entre imagens em (LIN et al., 2013) foi utilizado o método SIFT (*Scale-invariant feature transform*), este, pode ser considerado como uma espécie de detector de bordas/linhas que apresentam características padrões na imagem (LOWE, 2001). Segundo (MIKOLAJCZYK; SCHMID, 2005), o algoritmo SIFT é um dos métodos que obtêm os melhores resultados de correspondência entre pontos, levando em consideração os problemas mencionados.

Em cada uma das imagens, já com as distorções compensadas, o algoritmo é utilizado a

fim de identificar *key points* (também chamados de *pontos descritores*, em português), estes pontos são responsáveis pela identificação das bordas mais relevantes da imagem (normalmente baseados em uma frequência). Na Figura 9 é possível visualizar uma imagem com seus *key points* destacados em vermelho.

Em pares, as imagens são submetidas a um processo de comparação entre seus respectivos *key points*, identificando características em comum entre os pontos descritores de ambas. Os pontos descritores que tem a maior taxa de similaridade entre si (pixels vizinhos que apresentam o mesmo padrão, por exemplo) são considerados o mesmo ponto no espaço, ou seja, o ponto ideal de costura entre as imagens. Para a realização dessa tarefa, é necessário gerar uma matriz de homografia.

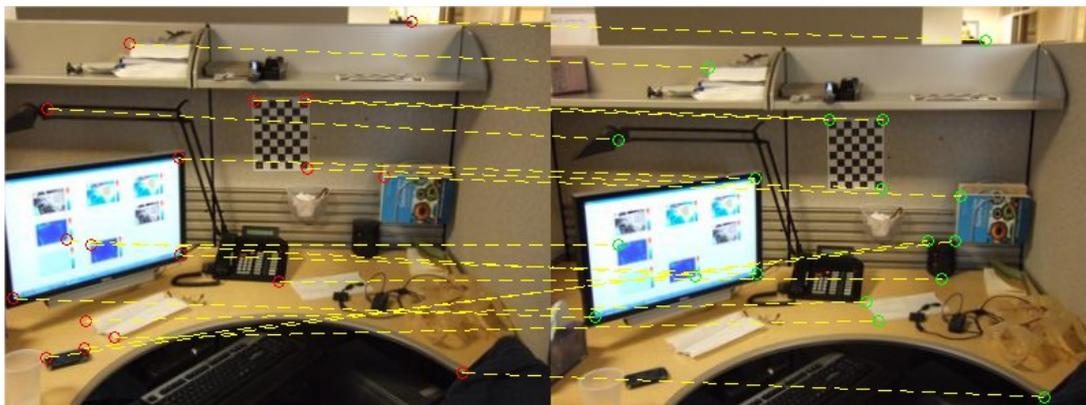
Figura 9 – Destaque dos *key points* da imagem, obtidos com a utilização do SIFT.



Fonte: Wang et al. (2017).

Em (GLEDHILL; TAYLOR; CLARKE, 2003) uma matriz de homografia é definida como um mapeamento entre duas imagens de uma superfície planar, constituindo uma única cena, adquiridas de perspectivas diferentes. Ou seja, duas imagens de uma mesma cena estão relacionadas por uma matriz de homografia (normalmente de dimensão  $3 \times 3$ , representando escala, rotação e translação entre as imagens). Essa matriz pode ser utilizada para remapear as imagens a partir de diferentes planos. Na Figura 10 é possível observar o resultado da correspondência de características (obtidas através do SIFT) entre imagens da mesma cena, obtidas de planos distintos.

Figura 10 – Exemplo de correspondência entre pontos com a utilização do SIFT.



Fonte: (MATHWORKS, 2017).

A fim de evitar problemas de registro, é necessário verificar se pontos de correspondência não foram estimados incorretamente. Para isso, é utilizado o algoritmo RANSAC (*Random Sample Consensus*), que é um método iterativo para estimar um modelo matemático de um conjunto de dados que contém pontos que fogem do padrão estatístico, chamados de *outliers*, e pontos que estão dentro deste padrão, denominados *inliers* (FISCHLER; BOLLES, 1981). Com a classificação desses dados concluída, é possível gerar uma nova matriz de homografia considerando apenas as coordenadas em ambas as imagens que estejam dentro do modelo matemático gerado. A remoção dos *outliers* é demonstrada na Figura 11, melhorando consideravelmente a taxa de acerto na correspondência entre pontos nas duas imagens.

Figura 11 – Remoção dos *outliers* com a utilização do RANSAC.



Fonte: (MATHWORKS, 2017).

Finalmente, com a melhor matriz de homografia definida, agora é possível realizar a fusão deste par de imagens. Em (LIN et al., 2013), os autores não mencionam o método que utilizaram para realizar a fusão das imagens, todavia é necessário utilizar algum método para corrigir artefatos que são causados pela sobreposição das imagens unidas. Realizando a união de todas as imagens, é possível visualizar todo o perímetro da garrafa com uma única imagem.

### 3 MÉTODO PROPOSTO

Este capítulo trata da estratégia utilizada para a aquisição das imagens que será utilizada na implementação do algoritmo. Além disso, é definido o escopo e são apresentados detalhes da implementação do método proposto em (LIN et al., 2013), que serve de base para o mapeamento das distorções avaliadas neste trabalho. Os métodos adotados para a compensação das distorções e montagem do mosaico também são descritos a seguir.

#### 3.1 PLATAFORMA DE AQUISIÇÃO

O experimento será realizado basicamente com três elementos: um objeto cilíndrico de referência (garrafa); uma câmera digital para obtenção das imagens em torno do eixo do objeto cilíndrico; e, por fim, uma plataforma que auxilia no movimento rotacional da garrafa.

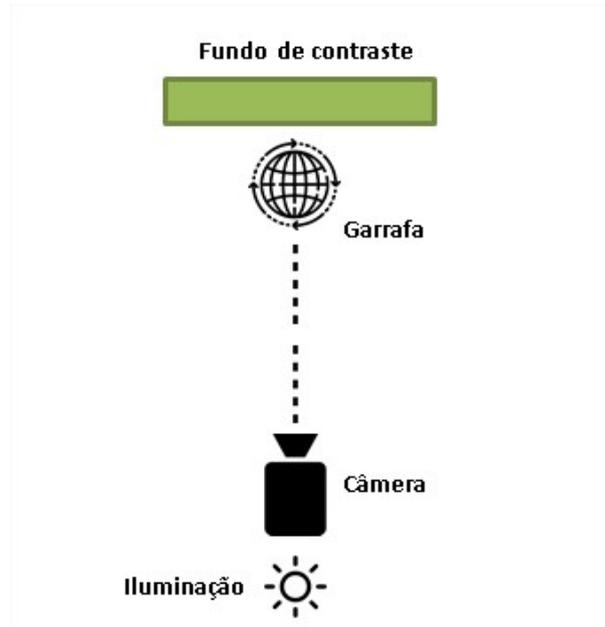
O objeto de interesse trata-se de uma garrafa, de vinho, com rótulo frontal e traseiro. Ainda que a produção de vinhos seja substancialmente menor que a produção de cerveja, uma garrafa típica de cerveja tem um formato mais próximo de um cilindro que a garrafa considerada, o que leva a menores distorções e, portanto, a um desempenho melhor dos algoritmos.

O dispositivo de aquisição utilizado é a câmera traseira de um *smartphone* ASUS, modelo Zenfone 3, com resolução de  $4656 \times 2620$  píxeis. A câmera é posicionada para aquisições em formato retrato (*portrait*) a uma distância adequada ao enquadramento integral da garrafa no campo de visão, adicionadas pequenas margens de segurança.

A garrafa e o sistema de aquisição são colocados em frente a um fundo verde, de forma a gerar uma imagem de alto contraste e de facilitar a segmentação, que não está propriamente no escopo deste trabalho. A câmera (única) é mantida fixa, enquanto a garrafa é rotacionada em passos de  $90^\circ$ . Com isso, são simuladas aquisições de imagens de uma garrafa fixa a partir de diferentes câmeras.

A rotação da garrafa é implementada com o auxílio de uma plataforma giratória. Essa plataforma possui um formato quadrado com a intenção de facilitar a rotação controlada de  $90^\circ$ , podendo ser dividida em duas partes, sendo a parte inferior responsável pela fixação e preservação do local em que a plataforma se encontra e a parte superior encarregada de realizar o movimento de rotação da garrafa. Os lados de ambas partes (superior e inferior) sempre permanecem alinhados, garantindo a movimentação correta entre a aquisição das imagens. Na plataforma superior são inseridos marcadores angulares, simulando a posição (angular) das câmeras dispostas em torno do eixo da garrafa. A Figura 12 apresenta um esquemático da plataforma utilizada.

Figura 12 – Diagrama esquemático da plataforma utilizada para a aquisição das imagens.

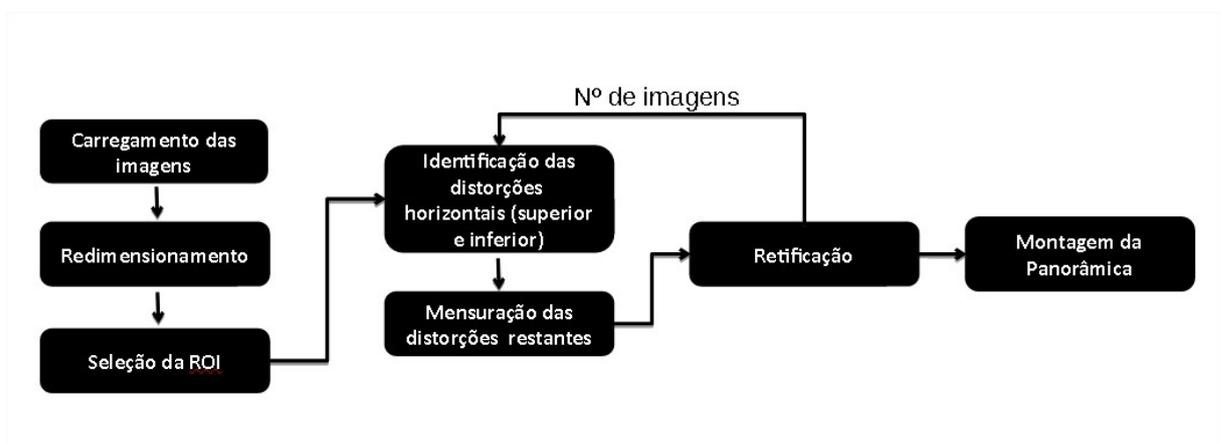


Fonte: O autor (2020).

### 3.2 SISTEMA PROPOSTO

Inicialmente, a imagem é submetida a um processo de redimensionamento e seleção da região de interesse. Após esse procedimento, é possível realizar a identificação das distorções de acordo com a quantidade de pontos escolhida pelo usuário. Realizando as compensações geométricas entre os pontos gerados pelo algoritmo, a união das quatro imagens é efetuada, possibilitando uma visão em torno de todo o eixo da garrafa. Um diagrama do sistema proposto é demonstrado na Figura 13.

Figura 13 – Etapas do sistema proposto.



Fonte: O autor (2020).

### 3.3 PRÉ-PROCESSAMENTO DA IMAGEM

As imagens, originalmente com  $4656 \times 2620$ , são redimensionadas para padrões HD ( $1280 \times 720$ ). Esse redimensionamento é feito por meio de interpolação bicúbica. São considerados dois tamanhos de imagens para possibilitar uma avaliação da perda de desempenho do método frente à redução da resolução. A redução de dimensão das imagens se dá em função do custo de implementação (o valor de câmeras industriais de resolução superior a HD tende a subir significativamente (WEBB, 2003; V B; DAVID, 2015)). Outra razão é a redução no custo computacional, o que, de forma semelhante, impacta nos custos de implementação de um sistema desse tipo.

Seis pontos de interesse, caracterizados pelos limites extremos do maior rótulo (quando observado frontalmente pela câmera) são definidos (por ora, de forma supervisionada). Os pontos que delimitam a região de interesse do algoritmo visam o dimensionamento da elipse causada pela distorção projetiva vertical e são caracterizados pelos seguintes limites (ver Figura 14):

- 1) Ponto superior esquerdo;
- 2) Ponto de maior ordenada;
- 3) Ponto superior direito;
- 4) Ponto inferior esquerdo;
- 5) Ponto de menor ordenada;
- 6) Ponto inferior direito.

Com as coordenadas desses pontos definidas, é possível observar a curva caracterizada geometricamente com o formato de um segmento de elipse, confirmando a representação da distorção afirmada por (LIN et al., 2013). Na Figura 14 é representado um exemplo da seleção de pontos em uma garrafa em conjunto de uma ilustração da elipse obtida.

Figura 14 – Pontos que representam as extremidades do rótulo (em amarelo) e a elipse imaginária (em vermelho).



Fonte: O autor (2020).

### 3.4 MAPEAMENTO DAS DISTORÇÕES

Para a identificação das distorções projetivas horizontais e verticais é necessário atribuir a quantidade de pontos horizontais que irão representar o “passo” de ângulo no plano da imagem. Uma quantidade de pontos maior proporciona um detalhamento mais fiel à imagem planificada, todavia isso interfere significativamente no desempenho do algoritmo, demandando um custo computacional mais elevado. Visto que o mapeamento das distorções é dado pelas equações (2.1) e (2.3), o algoritmo utilizado para a obtenção destas coordenadas é visto no pseudo-código descrito no Algoritmo 1.

A variação de  $\alpha$  permite obter as coordenadas que representam ambas as distorções. O que se pode perceber é uma forma geométrica de uma elipse (distorção vertical) com pontos horizontais menos afastados entre si conforme aproximam-se das bordas (limites do rótulo) (distorção horizontal).

Assumindo o diagrama 4, nota-se que a caracterização da distorção do eixo vertical vai causando uma perda da excentricidade da elipse conforme se aproxima do centro da figura. Com essa informação é possível concluir que novas elipses devem ser identificadas na área que se delimita aos pontos extremos já conhecidos.

Para realizar a identificação dos pontos internos (entre as duas elipses estabelecidas supervisionadamente), é possível estabelecer a relação entre as coordenadas das distorções já conhecidas através do pseudo-código transcrito no Algoritmo 2, *CES* e *CEI* correspondem à

---

**Algorithm 1** Mapeamento da Elipse

---

```

1: function MAP_ELLIPSE
2:    $i \leftarrow 1$ 
3:   for  $\alpha \leftarrow -\pi/2$  até  $\pi/2$  incremento de  $\pi/HP$  do  $\triangleright$  HP = Pontos horizontais
4:      $u \leftarrow$  Equação 2.1
5:     if elipse inferior then
6:        $v \leftarrow$  Equação 2.3
7:     else
8:        $v \leftarrow -$  Equação 2.3
9:     end if
10:     $Resultado(i) \leftarrow [u, v]$ 
11:     $i \leftarrow i + 1$ 
12:  end for
13:  return  $Resultado$ 
14: end function

```

---

coordenada do mesmo ângulo na elipse superior e na elipse inferior, respectivamente.

Basicamente a operação que o algoritmo realiza é a definição de um "passo" (que depende da quantidade de pontos verticais), responsável por diminuir a excentricidade da elipse superior. A elipse superior é escrita repetitivamente com a inclinação cada vez menor, até chegar próximo do centro (onde sua inclinação chega muito próxima de zero).

---

**Algorithm 2** Mapeamento das demais distorções

---

```

1: function DISTORTION_POINTS
2:   for  $indiceLinha \leftarrow 1$  até  $indiceLinha \leftarrow VP$  do  $\triangleright$  VP = Pontos verticais
3:     for  $indiceColuna \leftarrow 1$  até  $indiceColuna \leftarrow HP$  do
4:
5:        $delta \leftarrow \frac{CES[indiceColuna] - CEI[indiceColuna]}{VP}$ 
6:
7:        $row[indiceColuna] \leftarrow CES[indiceColuna] - delta * VP$ 
8:     end for
9:      $map[indiceLinha] \leftarrow row$ 
10:  end for
11:  return  $map$ 
12: end function

```

---

### 3.5 COMPENSAÇÕES DAS DISTORÇÕES

Com todos os pontos que representam as distorções existentes no plano da imagem identificados, se vê necessário realizar compensações geométricas a partir deste mapeamento. Para a realização desta tarefa, os valores são armazenados em vetores que representam coordenadas de todos os pontos horizontais a cada índice vertical. A partir deste vetor de armazenamento, sequencialmente, as coordenadas que representam um paralelogramo na imagem, são submetidas a uma transformação geométrica projetiva que irá alterar a forma geométrica (do então parale-

logramo), para um bloco com formato retangular com tamanho pré-definido (ver Figura 15).

Figura 15 – Aplicação da transformada geométrica projetiva no mapeamento das distorções.



Fonte: O autor (2020).

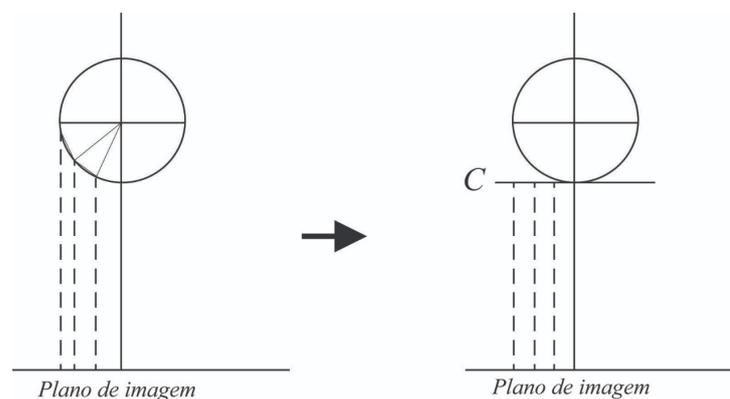
Visando uma representação fiel do rótulo analisado, o dimensionamento da imagem planificada é descrito a seguir. O dimensionado vertical da imagem (planificada) pode ser estabelecido através da distância entre pontos de maior e menor ordenada (definidos pelos pontos atribuídos supervisionadamente). Já o comprimento (total)  $C$  da imagem é dimensionado através do tamanho da circunferência que é visualizada na imagem. Utilizando a relação do perímetro da garrafa, é possível definir a largura total da imagem como sendo

$$C = \frac{2\pi r}{2} = \pi r \quad (3.1)$$

em que  $r$ , representa o raio da garrafa (em pixels).

Considerando que a imagem total é composta por blocos, cada um deles correspondendo a um ângulo  $\alpha$  no plano da imagem original, o dimensionamento horizontal  $C$  é dividido igualmente de acordo com a quantidade de pontos horizontais atribuídos. O diagrama da Figura 16 demonstra a lógica que estabelece a compensação geométrica entre os blocos.

Figura 16 – Esquemático do dimensionamento dos blocos que compõem a imagem.



Fonte: O autor (2020).

Realizando esse método de compensação com todos os respectivos pontos horizontais e verticais, a imagem planejada é finalmente concluída. Esse procedimento deve ser executado com as demais imagens que representam o deslocamento angular da garrafa, para assim, ser realizada a montagem do mosaico.

### 3.6 MONTAGEM DO MOSAICO

Diferente do método proposto em (LIN et al., 2013), que utiliza os algoritmos de registro mencionados no capítulo anterior, neste trabalho optou-se por assumir o movimento da garrafa como conhecido (ou seja, assume-se que a posição das câmeras é conhecida), para realizar a montagem da panorâmica. Os dois principais motivos que justificam a não utilização de algoritmos específicos para o registro de imagens são discutidos a seguir.

Em primeiro lugar, na abordagem de (LIN et al., 2013), o sistema é empregado para realizar a inspeção em garrafas PET (Polietileno Tereftalato), enquanto neste trabalho avalia-se a inspeção de garrafas que não possuam rótulo que cubra todo o perímetro da garrafa (diferente das garrafas PET, que em sua maioria, apresentam essa característica). Esse detalhe pode vir a impactar diretamente no registro das imagens, visando que as garrafas utilizadas neste trabalho são definidas por grandes áreas de cor homogênea, e a ausência de “bordas” e “cantos” pode vir a dificultar a detecção de *keypoints* na imagem. Por exemplo, se uma garrafa de vinho não apresentar um rótulo grande o suficiente para compartilhar áreas em comum entre as imagens, o registro pode não conseguir gerar *inliers* o suficiente para calcular uma matriz de homografia responsável por fazer o alinhamento correto das imagens.

O segundo fator aborda o custo computacional exigido pelo algoritmo de registro proposto pelo autor. O método SIFT é constantemente abordado na comunidade científica a fim de propor melhorias de desempenho ao algoritmo, que, apesar de apresentar bons resultados demanda um poder de processamento elevado para detectar os *keypoints* nas imagens (SUZUKI; IKENAGA, 2012). Somado a isso, na aplicação descrita nesse trabalho deve-se levar em conta o processamento da comparação entre as imagens (minimamente quatro), bem como a transformação e fusão, dificultando a aplicação do método em meio a sistemas de tempo real.

Considerando os problemas descritos, optou-se por assumir que os movimentos relativos entre as imagens são conhecidos. A caracterização do movimento conhecido estabelece que o algoritmo já conhece previamente a ordem de arranjo entre as cenas para a montagem do mosaico, bem como a representação angular de cada imagem na panorâmica.

Como já mencionado, o fato deste trabalho estar negligenciando calibrações intrínsecas à câmera e, por isso, cada uma das imagens planejadas é dimensionada para representar o valor ideal de meio perímetro da garrafa ( $180^\circ$ ), deve ser escolhida a porcentagem de visualização de cada uma das imagens para que cada segmento do mosaico não apresente blocos de imagens repetidos ou faltantes. Assumindo que serão obtidas quatro fotografias em torno do eixo da garrafa, sendo que cada uma dessas representará  $90^\circ$  de todo o seu perímetro  $C$  (idealmente),

optou-se pelo descarte de  $45^\circ$  em cada lado da representação planificada da imagem, assim, a soma da representação angular (na panorâmica) em cada uma das imagens, totalizará o perímetro completo da garrafa.

## 4 RESULTADOS

Este capítulo se destina à apresentação dos resultados obtidos diante da implementação do algoritmo. Na primeira seção, há uma breve descrição da plataforma utilizada para adquirir as imagens. Na segunda seção são demonstrados os resultados obtidos no mapeamento das distorções com a variação de alguns parâmetros fixos. Já na terceira seção, é possível analisar os resultados de imagens já retificadas e, por fim, a quarta seção demonstra o resultado do mosaico obtido após os processamentos realizados na imagem.

### 4.1 PLATAFORMA DE AQUISIÇÃO

As aquisições das imagens foram realizadas em uma plataforma dedicada a fim de obter a maior controle das condições de aquisição. A câmera utilizada teve seus parâmetros automáticos desativados e a exposição foi configurada manualmente, a fim de evitar variações de iluminação entre as diferentes imagens. A câmera fixada um tripé e teve sua inclinação ajustada com auxílio de uma ferramenta de nível. Na Figura 17 é possível visualizar o ambiente onde foram realizados os ensaios.

Figura 17 – Plataforma utilizada para a aquisição das imagens.



Fonte: O autor (2020).

## 4.2 IDENTIFICAÇÃO DAS DISTORÇÕES PROJETIVAS

Com a identificação supervisionada dos seis pontos que representam os limites do rótulo efetuada, é necessário fazer o mapeamento das distorções a partir dessas delimitações. As distorções foram identificadas a partir do Algoritmo 1 (Mapeamento da Elipse) com uma escolha de trinta pontos horizontais, facilitando a visualização da caracterização geométrica do método. Na primeira execução, os parâmetros fixos do sistema foram dados por  $R = 3,5$  e  $d = 33$  (ambos em cm). Na Figura 18 é possível observar os resultados diante desta reprodução, onde os pontos em vermelho descrevem o mapeamento das distorções e os pontos em amarelo, as coordenadas que delimitam o rótulo.

Figura 18 – Aplicação do algoritmo para obtenção da distorção.

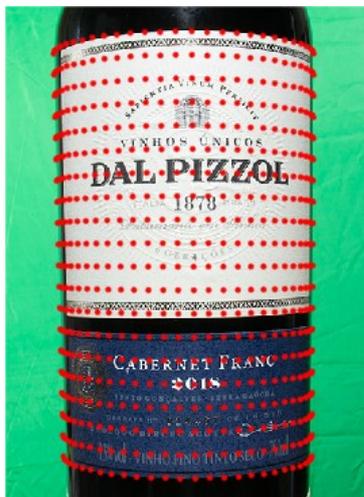


Fonte: O autor (2020).

O que se pode observar é que os pontos ficaram mais próximos entre si à medida que foram se aproximando do limite do rótulo (representando a distorção horizontal) e também atingiram um padrão geométrico esperado (elipse). As coordenadas restantes foram obtidas através do método descrito no capítulo anterior. Utilizando como parâmetro a presença de trinta pontos horizontais e vinte pontos verticais, a Figura 19 representa as distorções características a este rótulo perante o plano da imagem.

Nota-se que as distorções mapeadas apresentaram uma característica muito fiel ao diagrama esquemático da Figura 4, onde as elipses de maior ordenada apresentam uma inclinação superior às elipses que se concentram mais próximas do centro do rótulo. Também é possível destacar que o sistema de mapeamento proposto em (LIN et al., 2013) considera que o eixo focal da câmera se encontra minuciosamente alinhado com o centro do rótulo. A alteração de parâmetros fixos ( $R$  e  $d$ ) influencia diretamente nos quesitos excentricidade das elipses extremas e, em consequência disso, das elipses internas.

Figura 19 – Pontos que representam as distorções características desta garrafa.



Fonte: O autor (2020).

#### 4.3 REMOÇÃO DAS DISTORÇÕES

Para a validação desta etapa, a proposta de ensaio foi a utilização de uma garrafa envolvida por uma folha de papel quadriculado a fim de avaliar visualmente a retificação da imagem. Os desfechos a seguir demonstram o resultado da imagem planejada com diferentes quantidades de pontos verticais e pontos horizontais atribuídos. Na Figura 20 é possível observar uma fotografia do objeto utilizado para a avaliação de resultados. Todas as execuções foram realizadas com valores ( $N$ ) de pontos horizontais e verticais idênticos, proporcionando uma proporção de pontos de  $N \times N$ .

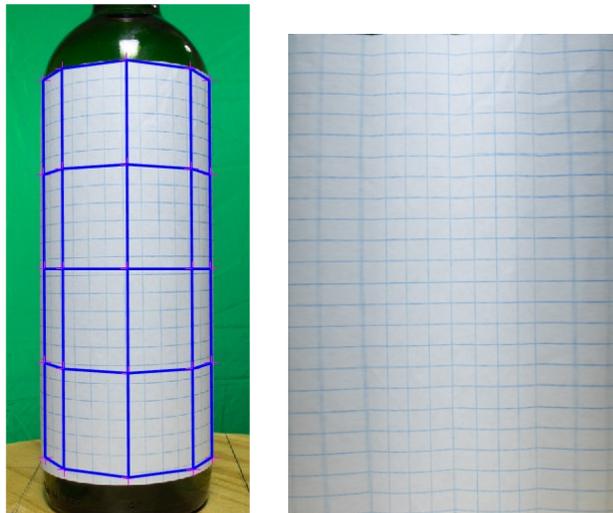
Figura 20 – Imagem do objeto analisado.



Fonte: O autor (2020).

A Figura 21 demonstra o primeiro experimento, no qual foi utilizado um valor de  $N = 5$ . A imagem à esquerda (a) representa a identificação das  $N$  distorções encontradas em cada eixo, enquanto a imagem à direita (b) representa o resultado da imagem planificada. Entre a Figura 22 e a Figura 24, os valores  $N$  são dados por 10, 20 e 30, respectivamente, e apresentam a mesma estrutura de apresentação. O que se pode perceber nestes resultados é a melhoria significativa da representação planificada da imagem conforme o valor de  $N$  é incrementado. No primeiro resultado, as falhas e distorções são facilmente visíveis conforme se aproximam do limite lateral da imagem. Contudo, os tamanhos dos retângulos são uniformes dentre a área mais central da imagem planificada.

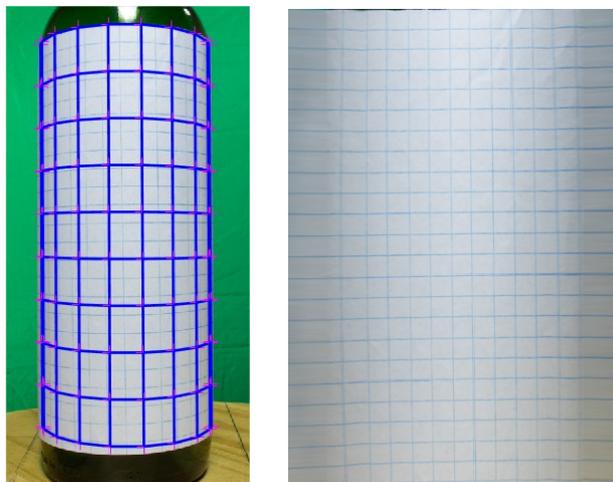
Figura 21 – Resultado da planificação com  $N = 5$ .



(a) Imagem original (b) Imagem planificada.

Fonte: O autor (2020).

Figura 22 – Resultado da planificação com  $N = 10$ .



(a) Imagem original (b) Imagem planificada.

Fonte: O autor (2020).

No segundo experimento, apesar de uma melhor atribuição ao tamanho dos retângulos, ainda é possível destacar a presença de borrões. Além disso, o último retângulo de cada lado não apresenta uma última linha vertical, formando um retângulo maior. Por fim, os últimos dois experimentos se demonstraram sem a presença de deformações aparentes dentre as imagens resultantes. Todavia, é importante observar que com  $N = 30$ , apesar de pouco perceptível, a imagem apresenta a última linha vertical levemente mais visível e uniforme, se comparada ao caso em que  $N = 20$ .

Figura 23 – Resultado da planificação com  $N = 20$ .

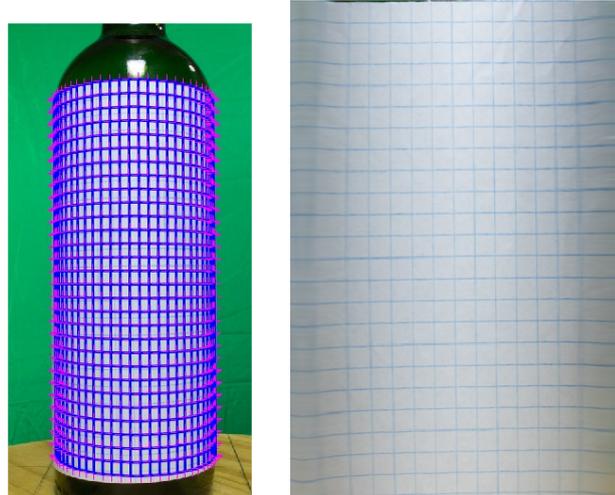


(a) Imagem original

(b) Imagem planificada.

Fonte: O autor (2020).

Figura 24 – Resultado da planificação com  $N = 30$ .



(a) Imagem original

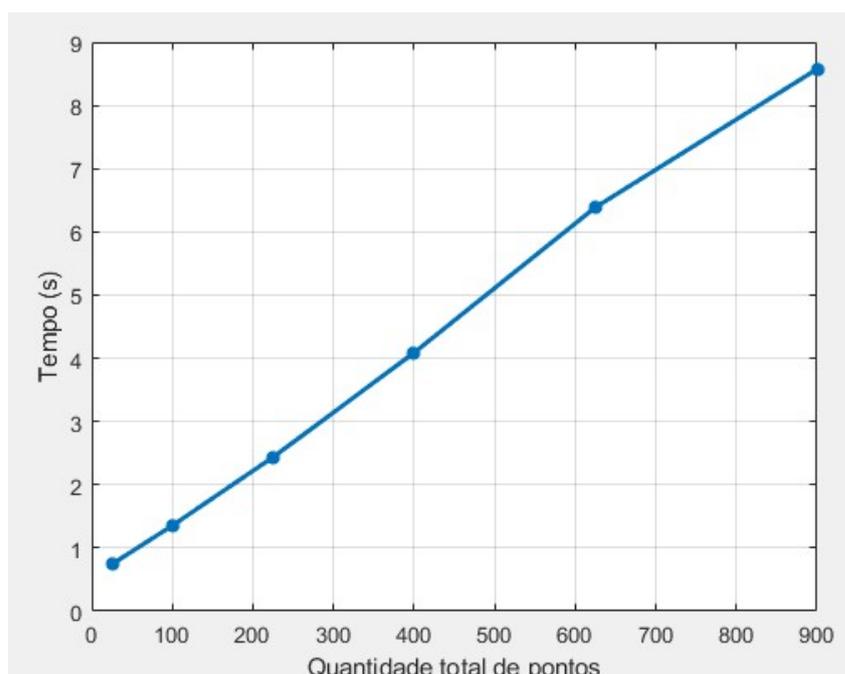
(b) Imagem planificada.

Fonte: O autor (2020).

Como mencionado, a utilização de uma maior quantidade de pontos de referência sobre a imagem resulta em um acréscimo significativo no custo computacional. Para realizar a medição

de tempo necessário para cada execução, foi utilizada a ferramenta *Profile* do MATLAB®, que retorna a demanda de tempo em segundos de acordo com a utilização de cada método executado no código. Os valores de tempo analisados se atribuem somente ao método de planificação da imagem, desconsiderando o tempo de redimensionamento da imagem, tempo de alocação, etc. No gráfico disposto na Figura 25 é demonstrado o tempo de execução do método de compensação das imagens de acordo com a quantidade de pontos.

Figura 25 – Tempo de execução (em segundos) para a quantidade de pontos total atribuída.



Fonte: O autor (2020).

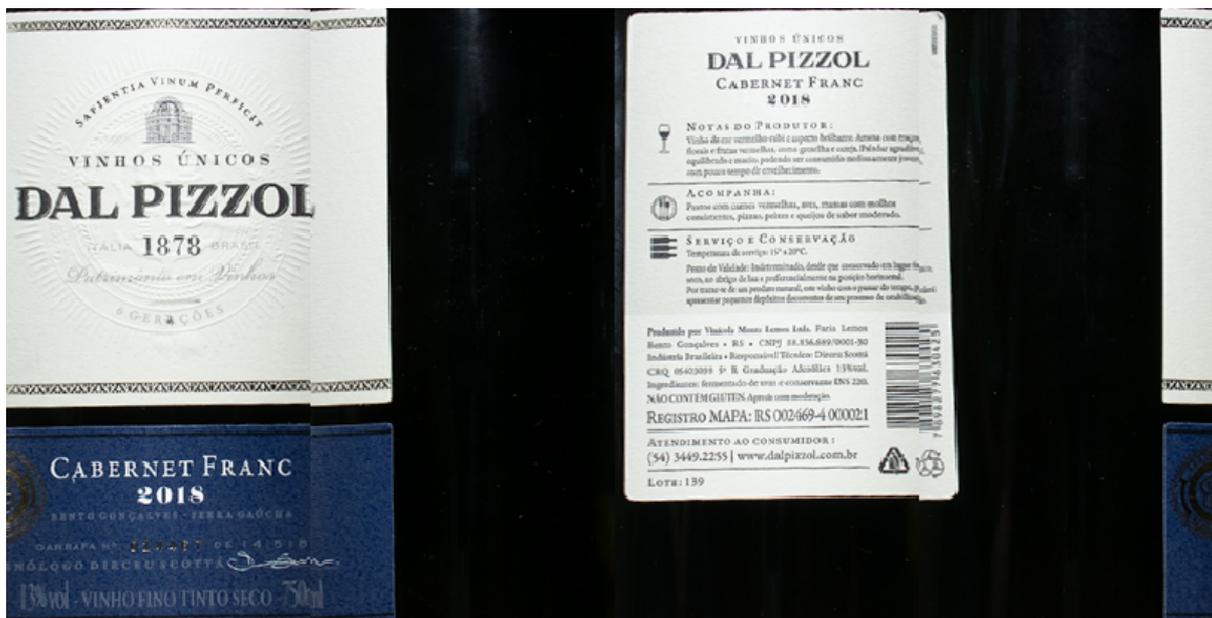
Observando-se o gráfico é possível afirmar que o tempo de execução cresce linearmente de acordo com o número de pontos atribuídos. A obtenção desses valores foi realizada seguindo as regras: execução de cinco realizações do algoritmo para cada valor de  $N$ , sendo considerada a média dos valores de cada realização como o valor utilizado no gráfico; em cada uma das execuções, o comando *clear all* era utilizado, a fim de evitar toda a pré-alocação de memória e outros fatores utilizados pela ferramenta para melhorar seu desempenho.

#### 4.4 MONTAGEM DA PANORÂMICA

Nesta seção é avaliada a montagem da panorâmica considerando o movimento conhecido da garrafa. Como mencionado, esse sistema considera que a cena é espacialmente estática, ou seja, a garrafa e a câmera sempre vão estar fixas, realizando apenas o movimento de rotação da garrafa. A Figura 26 demonstra o resultado da montagem da panorâmica formada a partir das figuras dispostas em Figura 27 (ver anexo A).

Cabe ressaltar a importância da precisão do sistema mecânico que compõe a solução pro-

Figura 26 – Resultado da panorâmica cortando 45°.



Fonte: O autor (2020).

posta sistema. Visto que o local de costura do mosaico é facilmente identificável pelos diferentes níveis de iluminação nas imagens, para comentar os detalhes, cada quarto de imagem será nomeado como A, B, C e D, respectivamente. Nota-se pelas regiões B e C que existe uma leve inclinação vertical no ponto de apoio da garrafa, o que interfere diretamente no alinhamento do sistema, que considera o posicionamento da garrafa como confiavelmente fixo, ou seja, quanto maior for o nível de controle do sistema mecânico, melhores podem ser os resultados obtidos através do método.

Neste resultado também é possível analisar que a desconsideração de 45° (1/4) em ambos os lados de cada imagem, apesar de terem sido utilizados valores teórico e ideais para o dimensionado aqui apresentado, resultou em uma boa representação de ponto de costura entre as imagens. Isto pode visto na junção do "L" ("em Pizzol"), bem como a extensão legível na parte inferior da costura entre os segmentos de imagem A e B.

Considerando que o presente trabalho conta com diversas "não-idealidades" para a implementação, o método conseguiu demonstrar através de uma única imagem todo o perímetro da garrafa, mesmo que com algumas falhas. Supondo um sistema mecânico superior em níveis de precisão, bem como o conhecimento dos principais parâmetros da câmera que está sendo utilizada, os resultados de montagem da panorâmica podem ser melhorados a fim de corrigir pequenas imperfeições.

## 5 CONSIDERAÇÕES FINAIS

Neste trabalho foi proposta uma etapa de um sistema computacional responsável por inspecionar rótulos de envazados com formato essencialmente cilíndricos. Aqui, se tratou do problema a partir do ponto de vista do processamento de imagens, com o propósito de corrigir as distorções causadas pelo formato do objeto quando reproduzido no plano do sensor da câmera, bem como a formação de uma panorâmica que abrange todo o perímetro do objeto.

Como o sistema proposto utiliza somente uma câmera física, emulando as outras três com translações do objeto (com passos angulares de  $90^\circ$ ) a fim de capturar a informação de todo o perímetro da garrafa, assume-se que o ponto no espaço ocupado pela garrafa é o mesmo nas quatro fotografias consideradas na composição da panorâmica.

As distorções provenientes ao formato do objeto (cilíndrico) são mapeadas através das equações (2.1) e (2.3) a partir da variação angular do sistema de coordenadas da imagem. O sistema negligencia distorções provenientes da lente da câmera, sendo esse um fator que pode ser implementado em trabalhos futuros. A calibração desse parâmetro permite um dimensionamento melhor da representação angular real de cada uma das fotografias, auxiliando na montagem do mosaico. Também é importante destacar que o mapeamento proposto por (LIN et al., 2013) exige uma confiança mecânica do sistema de aquisição que, apesar de ser minucioso, não é um grande problema para aplicações industriais.

As compensações geométricas do sistema são realizadas através da seleção de quatro coordenadas (formando um paralelogramo) provenientes ao mapeamento, sendo estes submetidos a uma transformação geométrica projetiva, mudando seu formato para um retângulo. A realização desta operação com todos os pontos do mapa forma uma nova imagem planificada do rótulo.

Por fim, as quatro imagens retificadas são submetidas ao processo de montagem do mosaico. Este, apesar de depender de um sistema mecânico robusto e confiável, se demonstrou minimamente eficiente a ponto de permitir um ponto ideal de costura aceitável para uma inspeção menos delicada.

Trabalhos futuros também podem ser realizados na direção de avaliar uma figura de mérito adequada para a comparação de diferentes panorâmicas. Resultados preliminares indicam que as altas frequências das panorâmicas obtidas podem ser suficientes para a avaliação do correto (ou não) posicionamento do rótulo da garrafa, através da comparação com um padrão.

## REFERÊNCIAS

- BOUGUET, J.-Y. **Camera Calibration Toolbox for Matlab** . [Online; accessed 02-November-2020], [http://www.vision.caltech.edu/bouguetj/calib\\_doc/index.html](http://www.vision.caltech.edu/bouguetj/calib_doc/index.html).
- CONNOR, J. J. O.; ROBERTSON, E. F. **August Ferdinand Möbius**. 1997.
- FISCHLER, R.; BOLLES, M. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. **Commun ACM**, [S.l.], v. 24, p. 619–638, 01 1981.
- GLEDHILL, D.; TAYLOR, D.; CLARKE, D. Panoramic imaging - a review. **Computers & Graphics**, [S.l.], v. 27, p. 435–445, 06 2003.
- GONZALEZ, R. C.; WOODS, R. E. **Digital image processing**. [S.l.: s.n.], 2000.
- GRACIAS, N. et al. Fast image blending using watersheds and graph cuts. **Image Vision Comput.**, [S.l.], v. 27, p. 597–607, 01 2009.
- HECKBERT, P. S. **Fundamentals of texture mapping and image warping**. [S.l.]: Citeseer, 1989.
- KOVACS, L. et al. Three-dimensional recording of the human face with a 3d laser scanner. **Journal of plastic, reconstructive & aesthetic surgery**, [S.l.], v. 59, n. 11, p. 1193–1202, 2006.
- LEE, J.-J.; KIM, G.-Y.; CHOI, H.-I. An image registration algorithm based on cylindrical prototype model. In: COMPUTATIONAL SCIENCE - ICCS 2001, 2001, Berlin, Heidelberg. **Anais...** Springer Berlin Heidelberg, 2001. p. 37–43.
- LEE, W.-S.; MAGNENAT-THALMANN, N. Fast head modeling for animation. **Image and Vision Computing**, [S.l.], v. 18, n. 4, p. 355–364, 2000.
- LEVIN, A. et al. Seamless image stitching in the gradient domain. In: 2004. **Anais...** [S.l.: s.n.], 2004. v. 3024.
- LIN, J. et al. Label inspection of approximate cylinder based on adverse cylinder panorama. In: OPTOELECTRONIC IMAGING AND PROCESSING TECHNOLOGY, 2013. **Anais...** [S.l.: s.n.], 2013. v. 9045, p. 90451Y.
- LOWE, D. Object recognition from local scale-invariant features. **Proceedings of the IEEE International Conference on Computer Vision**, [S.l.], v. 2, 01 2001.
- MATHWORKS. **Estimate fundamental matrix from corresponding points in stereo images**. [Online; accessed 02-November-2020], <https://www.mathworks.com/help/vision/ref/estimatefundamentalmatrix.html>.
- MIKOLAJCZYK, K.; SCHMID, C. A performance evaluation of local descriptors. **IEEE transactions on pattern analysis and machine intelligence**, [S.l.], v. 27, p. 1615–30, 11 2005.

- PARK, Y. et al. An ellipsoidal model for generating realistic 3d facial textures. **International Journal of Computer Applications in Technology**, [S.l.], v. 46, p. 36–44, 01 2013.
- SHEN, X. Approximate correction of length distortion for direct georeferencing in map projection frame. **Geoscience and Remote Sensing Letters, IEEE**, [S.l.], v. 10, p. 1419–1423, 11 2013.
- SHUM, H.-Y.; SZELISKI, R. Panoramic image mosaics. **International Journal of Computer Vision**, [S.l.], v. 36, p. 101–130, 04 2000.
- SUZUKI, T.; IKENAGA, T. Sift-based low complexity keypoint extraction and its real-time hardware implementation for full-hd video. In: THE 2012 ASIA PACIFIC SIGNAL AND INFORMATION PROCESSING ASSOCIATION ANNUAL SUMMIT AND CONFERENCE, 2012. **Proceedings...** [S.l.: s.n.], 2012. p. 1–6.
- V B, S.; DAVID, J. Significance of dimensionality reduction in image processing. **Signal & Image Processing : An International Journal**, [S.l.], v. 6, p. 27–42, 06 2015.
- WEBB, A. R. **Statistical pattern recognition**. [S.l.]: John Wiley & Sons, 2003.
- WOLBERG, G. **Digital image warping**. [S.l.]: IEEE computer society press Los Alamitos, CA, 1990. v. 10662.
- ZHAO, H. et al. Development of a coordinate transformation method for direct georeferencing in map projection frames. **International Journal of Photogrammetry and Remote Sensing**, [S.l.], v. 77, p. 94–103, 03 2013.

**ANEXO A – MONTAGEM DA PANORÂMICA**

Figura 27 – Imagens utilizadas para a formação da panorâmica.



(a)



(b)



(c)



(d)

Fonte: O autor, 2020.

