

Trabalho Final de Processamento de Imagens

Victor Requia¹, Lucas Martins Dallabeneta¹

¹Centro de Ciências Tecnológicas – Universidade do Estado de Santa Catarina (UDESC)

{victorrequia, lucasdallabeneta}@gmail.com

Resumo. *Este trabalho aborda três tópicos distintos no processamento de imagens: o uso da Transformada de Fourier para remover um reticulado em uma imagem, o rastreamento de objetos em sequências de imagens e a aplicação de técnicas de morfologia matemática para extrair um número específico de uma imagem binária. Cada tópico apresenta suas próprias técnicas e objetivos, contribuindo para uma compreensão abrangente e aprofundada do processamento de imagens e suas aplicações.*

Abstract. *This work addresses three distinct topics in image processing: the use of Fourier Transform to remove a grid pattern in an image, object tracking in image sequences, and the application of mathematical morphology techniques to extract a specific number from a binary image. Each topic presents its own techniques and objectives, contributing to a comprehensive and in-depth understanding of image processing and its applications.*

1. Introdução

Este trabalho tem como objetivo explorar diferentes aspectos do processamento de imagens e sua aplicação em diferentes contextos. Serão abordados três tópicos distintos, cada um com suas técnicas e objetivos específicos. No primeiro tópico, exploraremos a aplicação da morfologia matemática para a obtenção de uma imagem final a partir de uma imagem binária inicial. O objetivo específico será extrair somente o número 4 desta imagem utilizando operações de erosão e dilatação. Através da definição de uma semente inicial e a aplicação de dilatações sucessivas, seremos capazes de obter a imagem desejada. Discutiremos os valores escolhidos para o elemento estruturante e apresentaremos o algoritmo em Python utilizado para realizar as operações morfológicas. Por fim, faremos uma análise dos resultados obtidos e possíveis melhorias ou extensões para o trabalho.

No segundo tópico, iremos focar no processamento de imagens utilizando

a Transformada de Fourier, com o propósito de remover um reticulado presente em uma imagem. Serão utilizados filtros no domínio da frequência e métricas de avaliação para garantir a eficácia da remoção do reticulado, ao mesmo tempo em que preservamos as características importantes da cena.

No terceiro tópico, nosso foco será o rastreamento de objetos em sequências de imagens. Iniciaremos com a captura de um vídeo caseiro, onde um objeto se desloca sobre um fundo constante. Por meio da função `cvMatchTemplate` e diferentes métodos de casamento, buscaremos determinar o melhor método para rastrear o objeto. A partir disso, construiremos um vídeo de saída que demonstrará o objeto rastreado nas imagens, juntamente com um segundo vídeo destacando as detecções de casamento do template. A tabela com os valores de casamento fornecerá informações adicionais sobre os resultados obtidos

Ao abordar esses três temas, esperamos explorar diferentes técnicas e conceitos do processamento de imagens, fornecendo uma visão ampla sobre suas aplicações e possibilidades. Cada tópico apresentará seus próprios desafios e resultados, contribuindo para uma compreensão mais completa dessa área tão relevante e em constante evolução.

2. Etapa Experimental

Os experimentos foram realizados utilizando a linguagem python e as bibliotecas destacadas pelo professor. Todas as questões estão na mesma pasta, com os nomes de arquivos sugestivos das questões.

3. Análise de Resultados

Na aplicação sobre o domínio da frequência, apesar dos valores de SSIM e MSE mostraram que a similaridade não foi grande, visualmente, conseguimos notar uma melhora na imagem e aproximação da imagem das folhas original. É interessante notar que, pequenos ajustes no limiar da frequência, podem alterar de forma considerável o resultado final, principalmente nos pontos onde existe maior quantidade de pixel (centro da frequência). Também podemos notar que, quando aplicamos a máscara de Fourier no eixo vertical (figura 4), as linhas verticais da imagem reticulada, acabam ficando menos visíveis. Assim como, aplicando a máscara para o eixo horizontal (figura 6), as linhas horizontais ficam menos visíveis. É interessante notar que, na imagem da transformação final, aplicando a

máscara de Fourier no eixo vertical e horizontal, perdemos alguns detalhes da imagem original.

Para a morfologia, conseguimos um bom resultado após a dilatação, removendo o número 4 da imagem. Para o elemento estruturante, foi analisado visualmente para determinar a matriz.

Para o último exercício, com base nos resultados obtidos ao realizar o rastreamento de um objeto em vídeos utilizando o método de template matching, foi possível concluir que o desempenho do algoritmo varia de acordo com o cenário em que é aplicado

4.1 Questão A

A Transformação de Fourier é uma técnica amplamente utilizada no processamento de imagens para analisar e manipular frequências presentes em uma imagem. Ela é baseada na ideia de decompor uma imagem em componentes sinusoidais que representam diferentes frequências.

O processo de Transformação de Fourier envolve os seguintes passos: Conversão da imagem para o domínio da frequência, Centro da frequência, Análise e manipulação das frequências, Modificação da imagem, Reconstrução da imagem.

A Transformação de Fourier é amplamente utilizada em várias aplicações de processamento de imagens, incluindo filtragem de imagens, compressão de imagens, análise de texturas, detecção de bordas e segmentação de imagens. Ela permite uma representação compacta das características de frequência presentes em uma imagem, facilitando a análise e a manipulação dessas características.

4.1.1 Máscara de Fourier aplicado no eixo vertical



Figura 1: Imagem Original

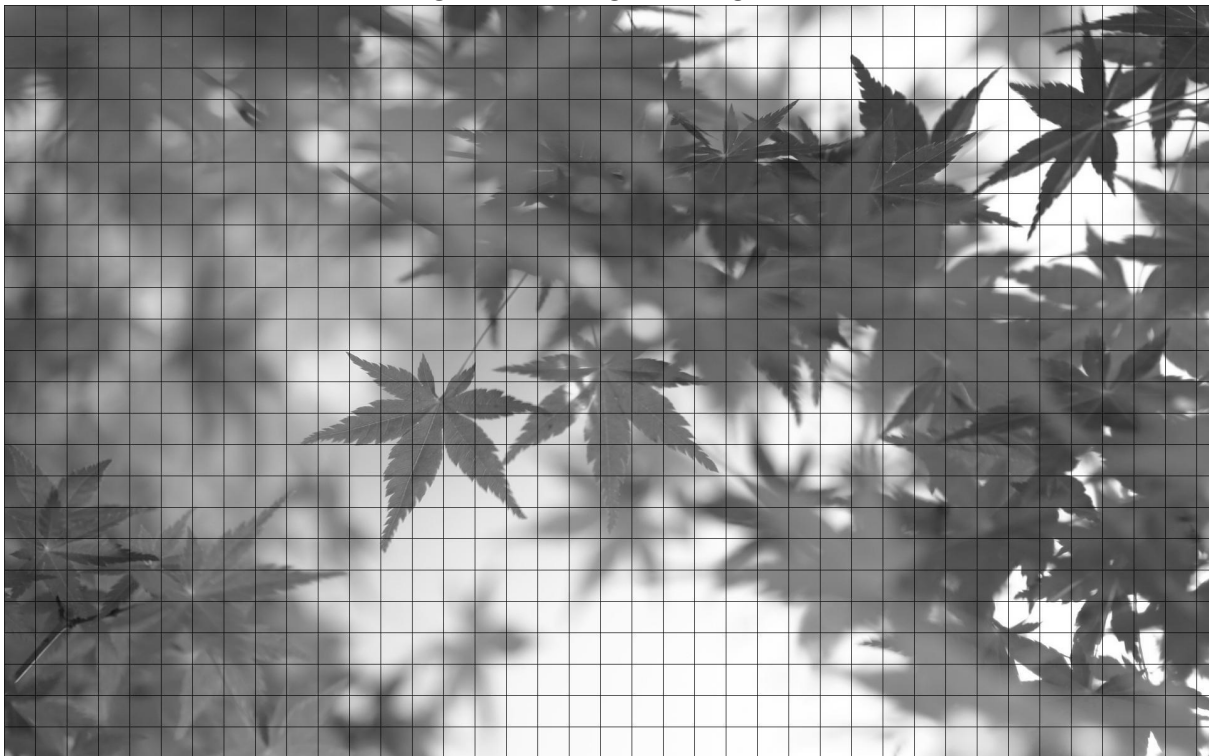


Figura 2: Imagem Original Reticulada

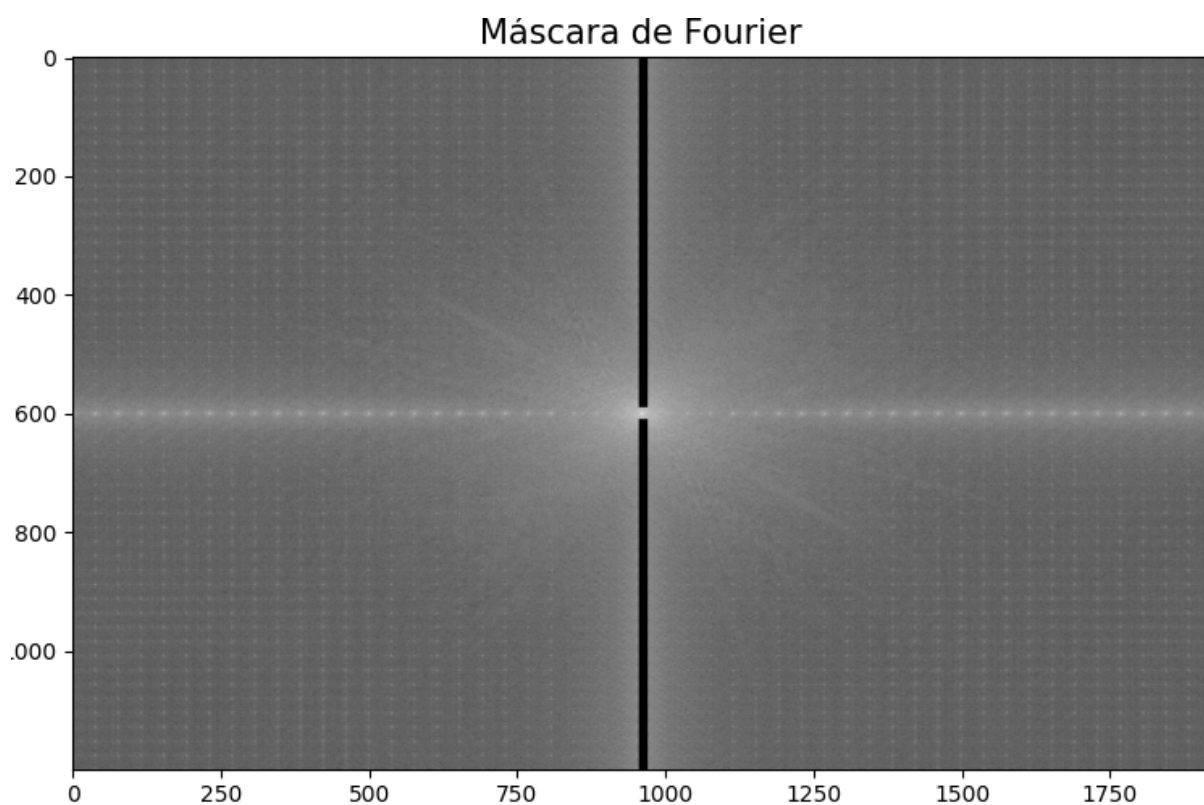


Figura 3: Domínio da frequência no eixo vertical

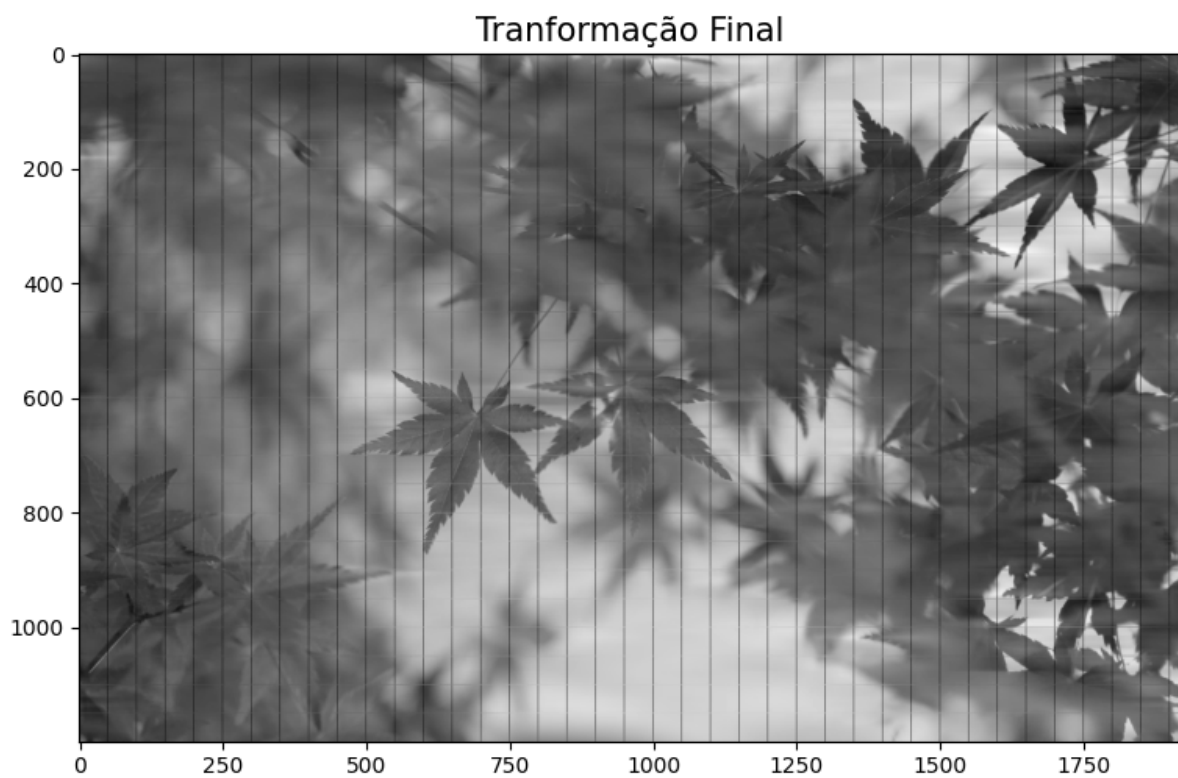


Figura 4: Aplicação da máscara de fourier no eixo vertical

4.1.2 Máscara de Fourier aplicado no eixo horizontal

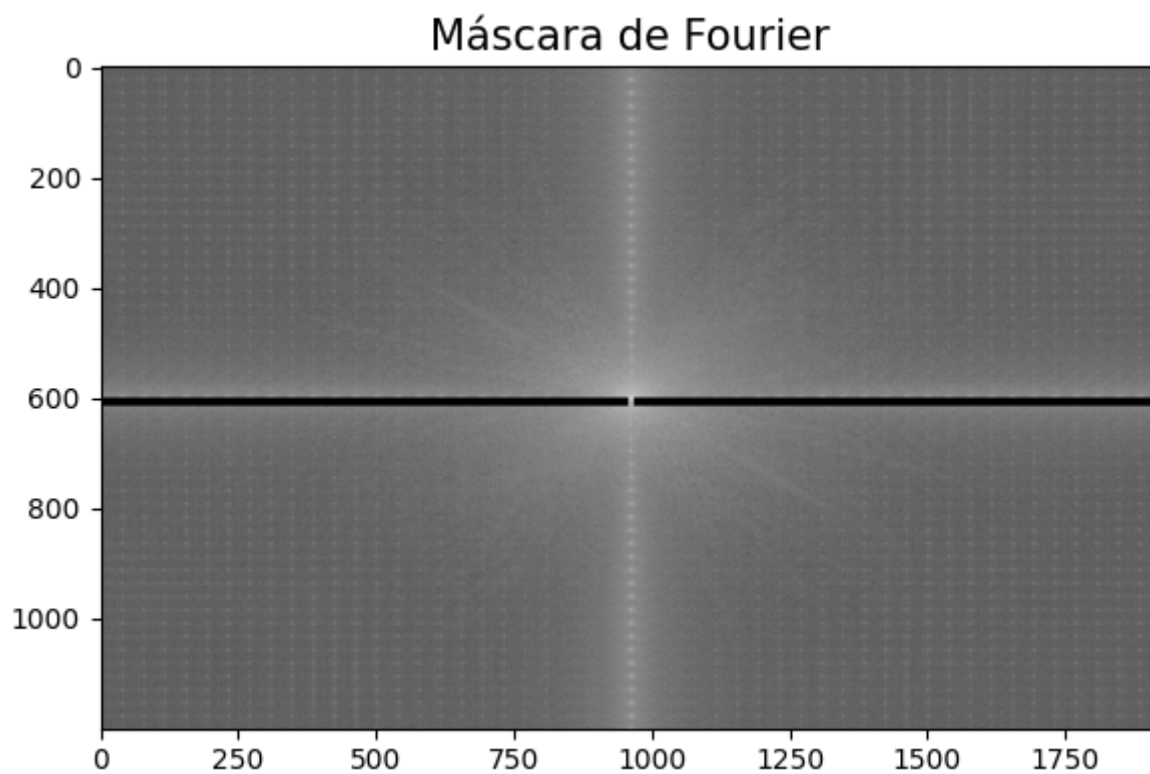


Figura 5: Domínio da frequência no eixo horizontal

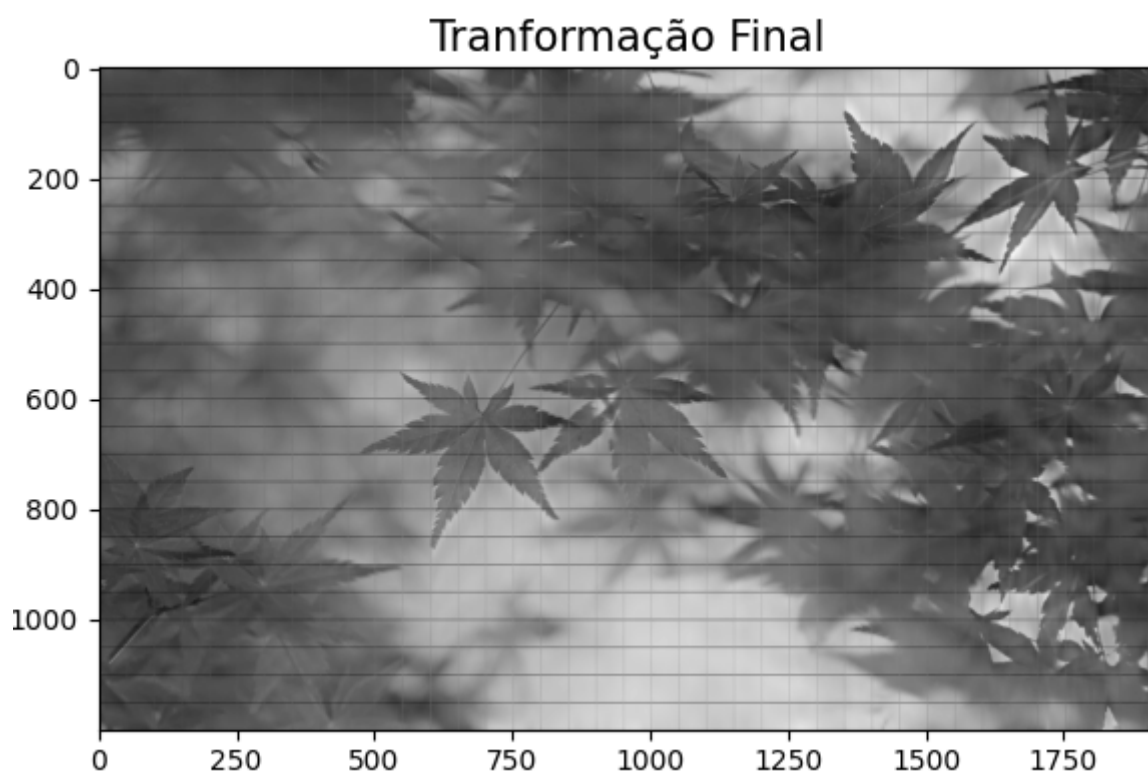


Figura 6: Aplicação da máscara de fourier no eixo horizontal

4.1.3 Máscara de Fourier aplicado no eixo vertical e horizontal

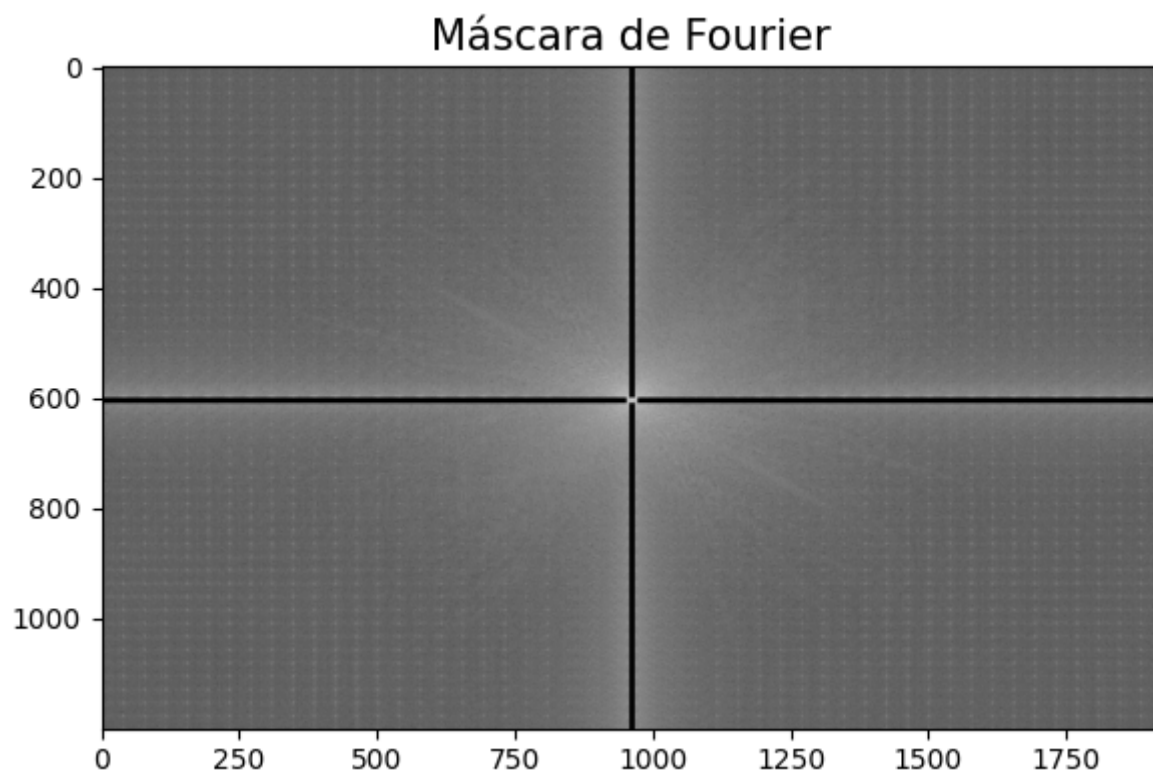


Figura 7: Domínio da frequência no eixo vertical e horizontal

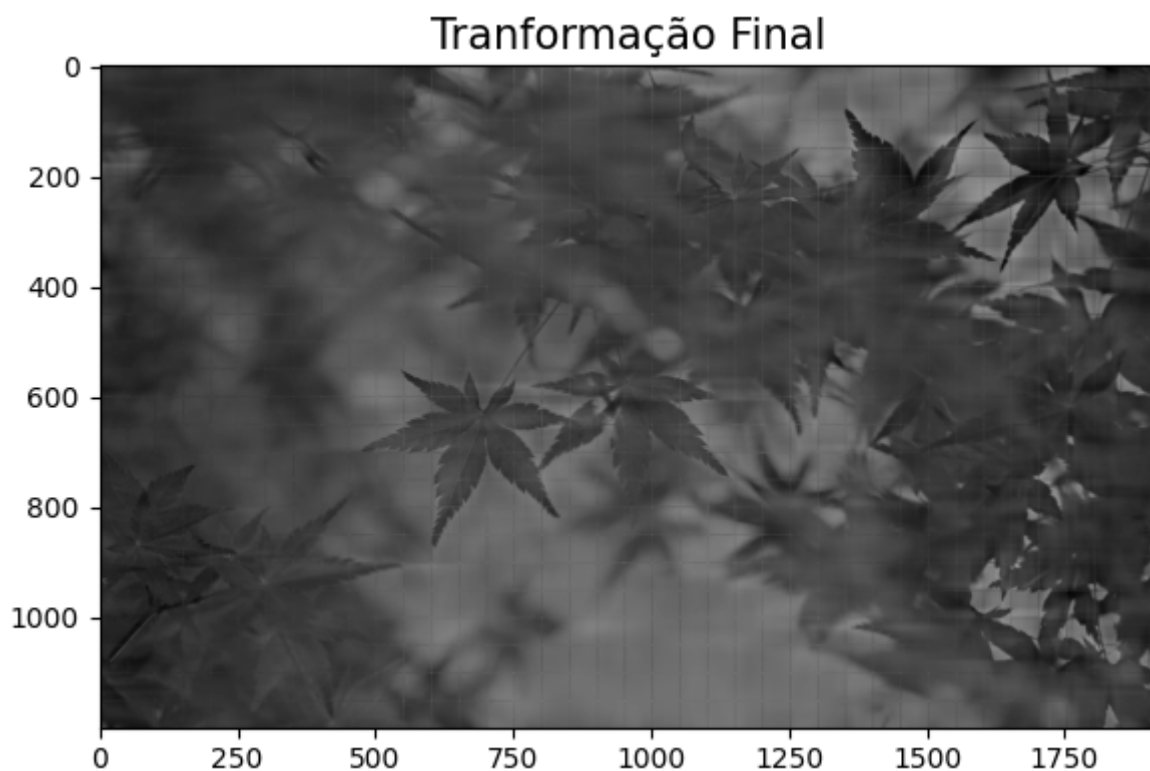


Figura 8: Aplicação da máscara de fourier no eixo vertical e horizontal

4.1.4 Comparação SSIM e MSE

O SSIM (Structural Similarity Index) e o MSE (Mean Squared Error) são medidas de avaliação da semelhança entre duas imagens. Essas métricas são amplamente utilizadas na área de processamento de imagem para comparar a qualidade de uma imagem modificada em relação à imagem original.

```
SSIM VERTICAL: 0.00927427024821278
MSE VERTICAL: 9458.524741319445
SSIM HORIZONTAL: 0.010362803109874348
MSE HORIZONTAL: 9464.312227430555
SSIM VERTICAL E HORIZONTAL: 0.010204847134285872
MSE VERTICAL E HORIZONTAL: 9489.809038628473
```

Figura 9: Comparação SSIM e MSE nos diferentes eixos da máscara de Fourier

SSIM VERTICAL: 0.00927427024821278

MSE VERTICAL: 9458.524741319445

Esses valores indicam que a transformação vertical aplicada à imagem resultou em um SSIM de 0.00927427024821278, indicando uma baixa similaridade estrutural em relação à imagem original. O valor do MSE foi calculado como 9458.524741319445, que representa a média dos erros quadrados entre os pixels da imagem original e os pixels da imagem transformada.

SSIM HORIZONTAL: 0.010362803109874348

MSE HORIZONTAL: 9464.312227430555

Nesse caso, a transformação horizontal aplicada à imagem resultou em um SSIM de 0.010362803109874348, também indicando uma baixa similaridade estrutural em relação à imagem original. O valor do MSE foi calculado como 9464.312227430555.

SSIM VERTICAL E HORIZONTAL: 0.010204847134285872

MSE VERTICAL E HORIZONTAL: 9489.809038628473

Aqui, a transformação vertical e horizontal aplicada à imagem resultou em um SSIM de 0.010204847134285872, mais uma vez indicando uma baixa similaridade estrutural em relação à imagem original. O valor do MSE foi calculado como 9489.809038628473.

Em geral, quanto mais próximo o valor de SSIM estiver de 1, maior será a similaridade entre as imagens. Um valor de SSIM próximo de 0 indica uma baixa similaridade estrutural. Já o MSE mede a média dos erros quadrados entre os pixels das imagens, onde um valor mais baixo indica uma menor diferença entre as imagens.

Com base nos valores, podemos concluir que as transformações aplicadas resultaram em uma baixa similaridade estrutural em relação à imagem original, pois os valores de SSIM são baixos. Além disso, o MSE indica uma diferença considerável entre os pixels das imagens original e transformada.

4.2 Questão B

Dilatação e erosão são operações duais em relação à complementação e reflexão de conjuntos:

$$(A \ominus B)^c = A^c \oplus \hat{B}$$

A erosão de A por B é o conjunto de todos os pontos x tais que B, quando transladado por x fique contido em A. Como no caso da dilatação, essa equação não é a única definição da erosão.

A erosão diminui ou afina os objetos em uma imagem binária. De fato, podemos ver a erosão como uma operação de filtragem morfológica em que os detalhes da imagem menores que o elemento estruturante são filtrados (removidos) da imagem.

Aqui estão alguns exemplos de matriz do elemento estruturante e seus resultados:

$B = [0, 0, 0],$
 $[0, 1, 0],$
 $[0, 0, 0]]$



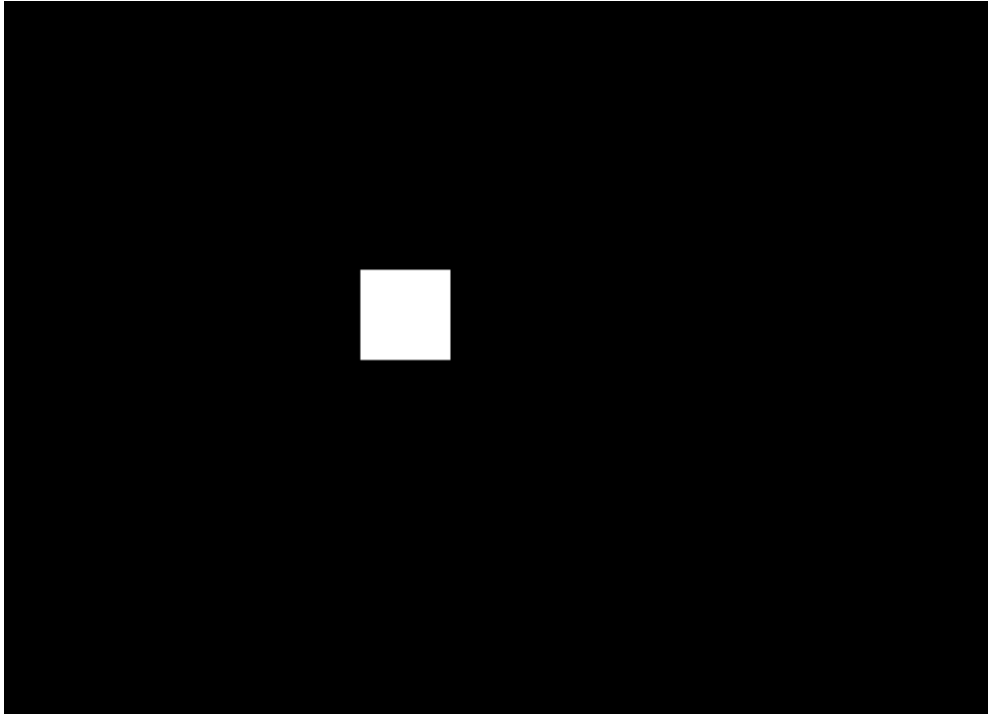
$B = [0, 0, 0],$
 $[0, 0, 0],$
 $[0, 0, 0]]$



$B = [0, 1, 0],$

$[1, 1, 0],$

$[0, 1, 0]]$



Com essa última matriz, vamos fazer a dilatação

A dilatação é uma operação morfológica utilizada no processamento de imagens para ampliar ou expandir áreas de interesse ou objetos presentes em uma imagem. Ela faz parte de um conjunto de técnicas conhecidas como transformações morfológicas, que manipulam a forma e a estrutura de objetos em uma imagem.

A dilatação é comumente usada para preencher lacunas ou buracos em objetos, conectar áreas próximas ou expandir os contornos dos objetos. É especialmente útil para melhorar a qualidade da segmentação de objetos e extrair características relevantes em uma imagem.

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\}$$

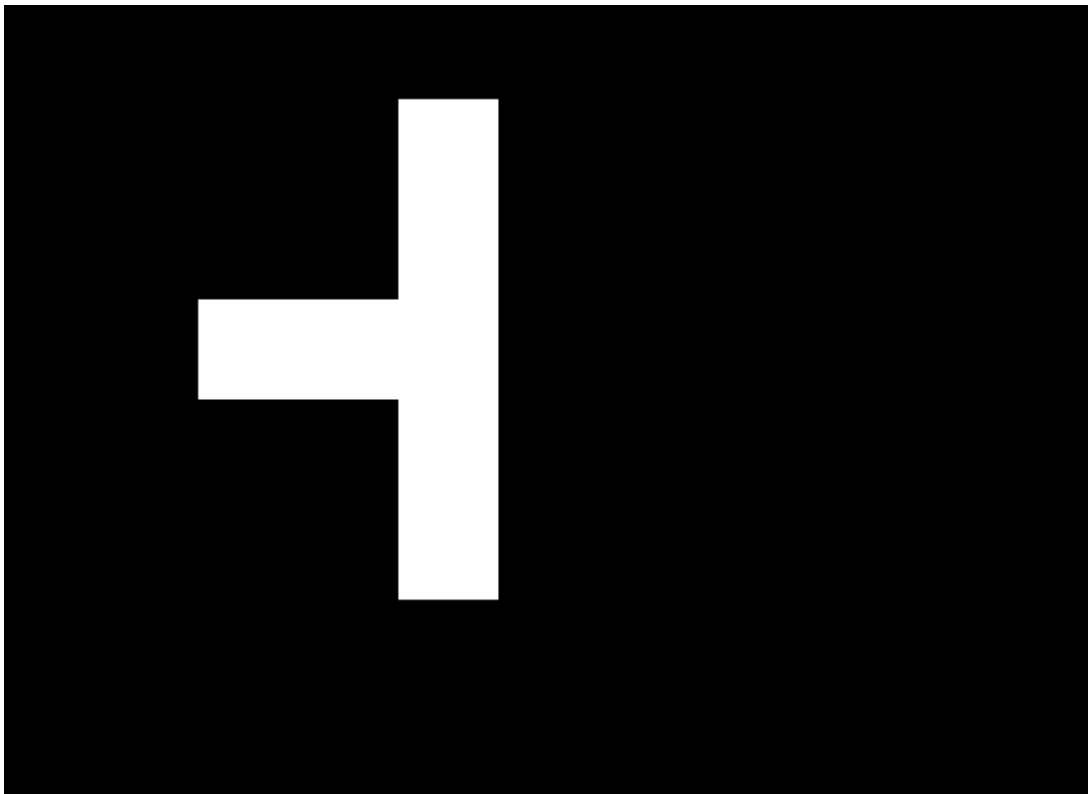


Figura 10: Iteração de dilatação = 2

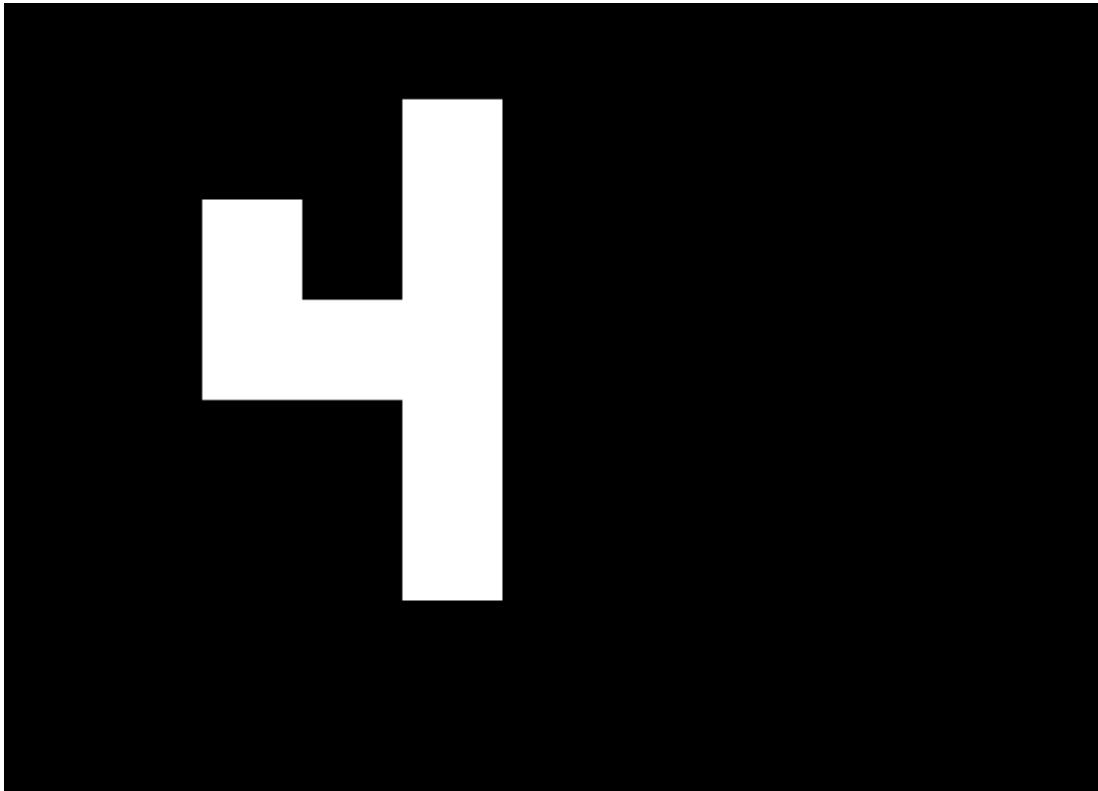


Figura 11: Iteração de dilatação = 3



Figura 12: Iteração de dilatação = 4

4.3 Questão C

A etapa C da tarefa consiste em realizar o rastreamento de um objeto em um vídeo utilizando o método de template matching. O template matching envolve comparar um padrão com uma imagem principal para identificar regiões semelhantes.

Para realizar o rastreamento, foi gravado um vídeo de uma bola se movendo em um cenário com fundo estático. Utilizando a biblioteca OpenCV, foram aplicadas as funções `cvMatchTemplate` e `cvMinMaxLoc`, que são especialmente projetadas para implementar o template matching. Essas funções permitem avaliar a similaridade entre o template e a imagem principal.

Para testar o método, foram utilizados diferentes algoritmos de casamento de template: `cv.TM_CCOEFF`, `cv.TM_CCOEFF_NORMED`, `cv.TM_CCORR`, `cv.TM_CCORR_NORMED`, `cv.TM_SQDIFF` e `cv.TM_SQDIFF_NORMED`. Cada algoritmo tem suas características e diferentes formas de calcular a similaridade.

Além de testar o rastreamento em um cenário com fundo estático, também foi realizado o teste em um ambiente mais complexo, com vários elementos visuais no fundo. Isso permite verificar como o algoritmo se comporta em situações mais desafiadoras.

Através da comparação dos resultados obtidos com cada algoritmo, é possível identificar qual deles apresenta o melhor desempenho. Isso pode ser avaliado considerando critérios como precisão e capacidade de lidar com variações no ambiente. Com base nos resultados, é possível ver algoritmos que se dão melhor com o vídeo gravado. Abaixo estão os resultados dos algoritmos citados anteriormente

cv.TM_CCOEFF

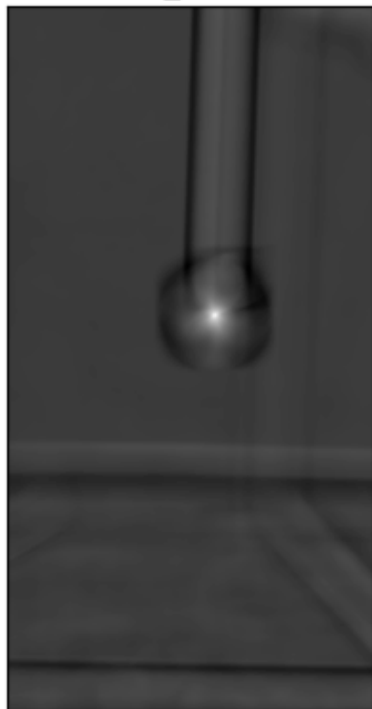


Figura 13: Método `TM_SQDIFF_TM_CCOEFF_NORMED` aplicado ao frame 336 do vídeo.

cv.TM_CCOEFF_NORMED

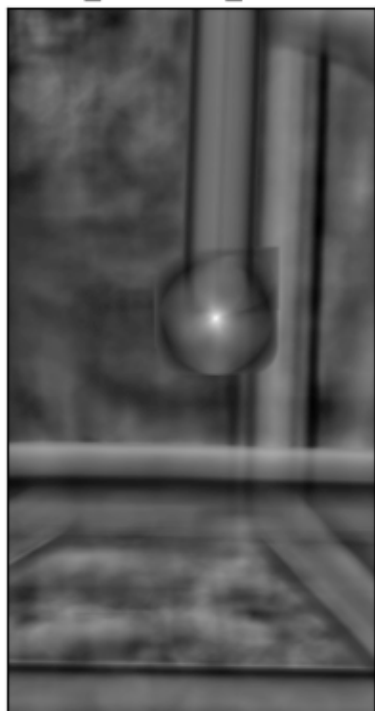


Figura 14: Método TM_CCOEFF_NORMED aplicado ao frame 336 do vídeo.

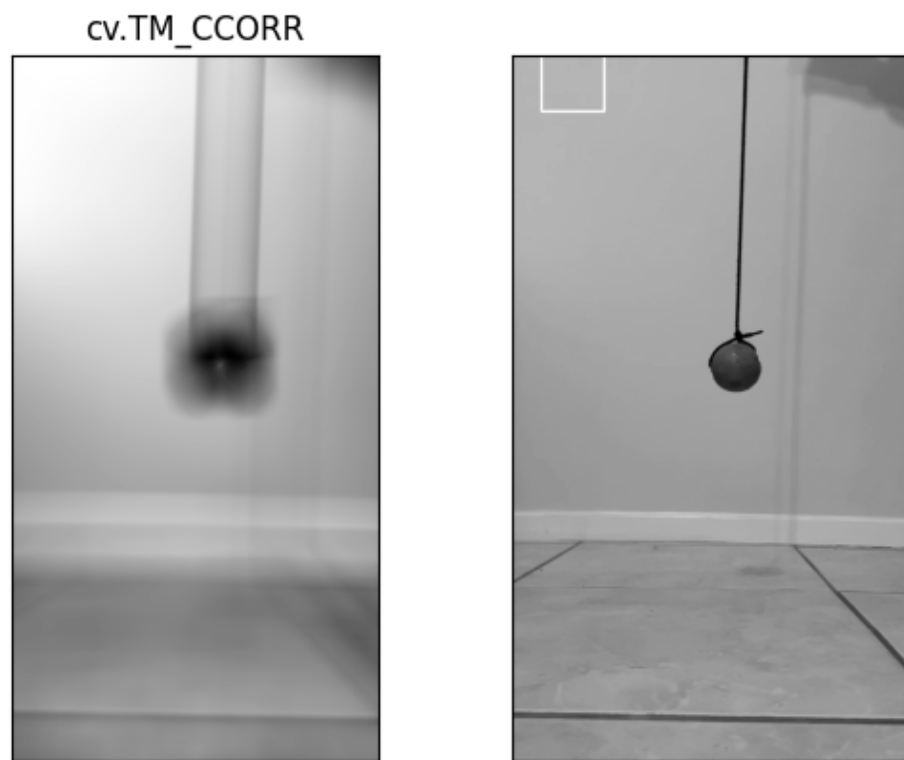


Figura 15: Método TM_CCORR aplicado ao frame 336 do vídeo.

cv.TM_CCORR_NORMED



Figura 16: Método TM_CCORR_NORMED aplicado ao frame 336 do vídeo.

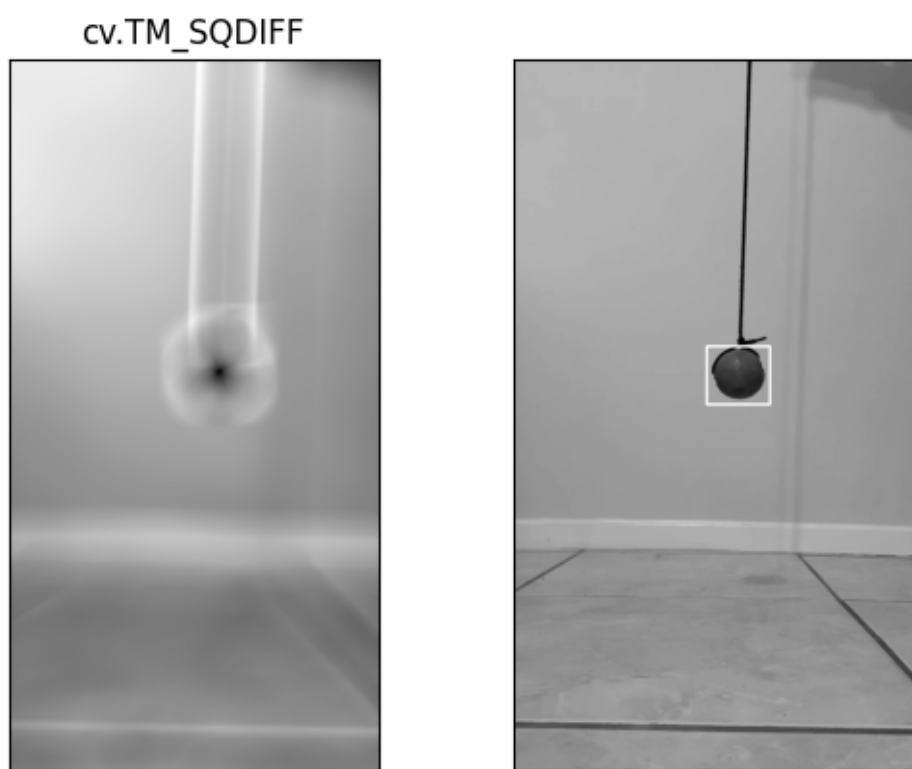


Figura 17: Método TM_SQDIFF aplicado ao frame 336 do vídeo.

cv.TM_SQDIFF_NORMED

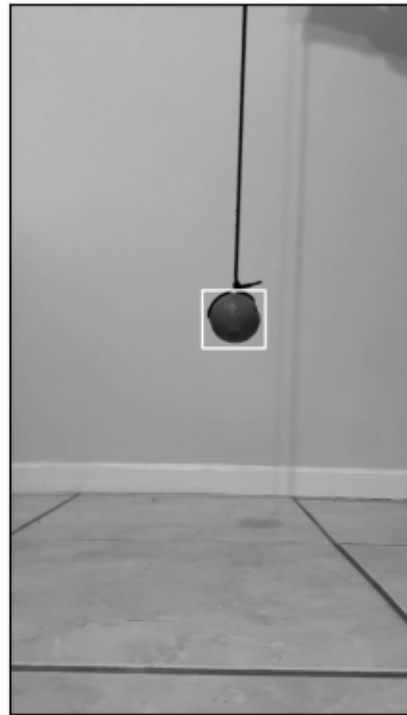


Figura 18: Método TM_SQDIFF_NORMED aplicado ao frame 336 do vídeo.

O algoritmo TM_CCORR (Template Matching Correlation) não conseguiu obter a posição da bola na imagem, o que talvez indique que o uso de correlação não seja a melhor alternativa nesse vídeo em específico.

Método	Mínimo	Máximo
cv.TM_CCOEFF	(199, 289)	(229, 343)
cv.TM_CCOEFF_NORMED	(343, 194)	(229, 343)
cv.TM_CCORR	(241, 333)	(35, 0)

cv.TM_CCORR_NORMED	(267, 334)	(220, 343)
cv.TM_SQDIFF	(229, 343)	(207, 1)
cv.TM_SQDIFF_NORMED	(229, 343)	(268, 334)

Tabela 1: Comparação entre os diferentes métodos usados

Os métodos cv.TM_CCOEFF e cv.TM_CCOEFF_NORMED apresentaram valores mínimos e máximos relativamente próximos entre si, o que sugere uma consistência em encontrar correspondências entre o template e a imagem principal. Ambos os métodos podem ser considerados como boas opções para o template matching nesta imagem. Por outro lado, o método cv.TM_CCORR mostrou uma faixa mais ampla de valores mínimos e máximos, o que indica uma maior variabilidade. Isso sugere que o método teve dificuldade em encontrar correspondências precisas em algumas partes da imagem.

4. Conclusão

Para a transformação de Fourier, podemos notar uma melhora visual significativa no que diz respeito às linhas reticuladas da imagem, porém os resultados visuais não foram os mesmos que os resultados da comparação com a imagem original utilizando SSIM e MSE. Nestes casos, os valores de similaridade foram pequenos, mostrando que as figuras não possuem uma semelhança na estrutura.

Para a morfologia matemática, conseguimos realizar o exercício com o resultado esperado. Conseguimos remover o número 4 apenas depois de aplicar a 4 iteração.

Com base nos resultados obtidos ao realizar o rastreamento de um objeto em vídeos utilizando o método de template matching, foi possível concluir que o desempenho do algoritmo varia de acordo com o cenário em que é aplicado. Quando foi utilizado um vídeo com fundo estático, os diferentes algoritmos de template matching conseguiram realizar o matching

com sucesso, fornecendo informações precisas sobre a posição da bola na imagem. No entanto, ao testar o algoritmo em um ambiente mais complexo, com vários elementos visuais no fundo, o resultado não foi satisfatório. Especificamente, o algoritmo TM_CCORR não conseguiu obter a posição correta da bola na imagem. Isso indica que o uso da correlação não é a melhor alternativa nesse vídeo específico. Em cenários mais complexos, pode ser necessário explorar outros métodos ou técnicas mais avançadas como o machine learning, para obter resultados mais precisos e robustos.