

VITOR HUGO FAZOLI DA SILVA

DESENVOLVIMENTO DE FERRAMENTA PARA
ILUMINAÇÃO EM IMAGENS BIDIMENSIONAIS DE
MÉDIA E BAIXA RESOLUÇÃO

Relatório Técnico elaborado conforme a
ABNT NBR 10719:10, apresentado ao
Instituto Federal de Educação, Ciência
e Tecnologia de São Paulo, como
parte dos requisitos para a obtenção
do grau de Bacharel em Ciência da
Computação.

Área de Concentração: Compu-
tação Gráfica

Orientador: Prof Dr. Gabriel Marcelino
Alves

SÃO JOÃO DA BOA VISTA

2023

SUMÁRIO

1	INTRODUÇÃO	5
1.1	Objetivos	6
1.1.1	Objetivo Geral	6
1.1.2	Objetivos Específicos	6
2	REVISÃO DA LITERATURA	7
2.1	Modelos de Iluminação	7
2.1.1	Reflexão Ambiente	7
2.1.2	Reflexão Difusa	8
2.1.3	Reflexão Especular	9
2.2	Sombreamento	9
2.2.1	Flat Shading	9
2.2.2	Sombreamento de Gourand	10
2.2.3	Sombreamento de Phong	10
2.3	Segment Anything em Visão Computacional	11
2.3.1	Codificador de Imagens (Image Encoder)	11
2.3.2	Codificador de Prompts (Prompt Encoder)	12
2.3.3	Decodificador de Máscaras (Lightweight Mask Decoder)	12
2.4	Geração de Mapas de Normais	13
2.4.1	Métodos Tradicionais	13
2.5	Técnicas avançadas de pixel arte	14
2.6	Ferramentas para iluminação de cenários	14
2.6.1	Vulkan	15
2.6.2	OpenGL	15
2.6.3	Engines	15
2.6.3.1	Unity	16
2.6.3.2	Godot	16
2.7	Trabalhos Correlatos	16
2.7.1	Geração De Mapas De Iluminação Baseado Em Topologia Estimada Para Iluminação De Sprites	17
2.7.2	Analysis and Compilation of Normal Map Generation Techniques for Pixel Art	17
2.7.3	CrossSketch: freeform surface modeling with details	17
3	METODOLOGIA	19
3.1	Levantamento de Requisitos	19

3.2	Estudo: Iluminação e Unity	20
3.3	Avaliar vantagens e limitações das ferramentas abordadas	21
3.4	Modelagem da ferramenta	21
3.5	Justificativa das Tecnologias a serem adotadas	22
3.6	Avaliação Qualitativa	23
4	ANÁLISE DOS RESULTADOS PRELIMINARES	24
4.1	Cronograma do Trabalho	24
5	CONCLUSÕES PARCIAIS	25
	REFERÊNCIAS	26

RESUMO

A iluminação é de grande importância para pinturas e esboços digitais nos dias de hoje, pois fornece a percepção de volume ao ambiente. Porém, a iluminação ainda é difícil de replicar em uma imagem digital, pois representa algo muito diferente de um ambiente tridimensional, ainda mais se usarmos uma imagem em baixa resolução. Nesse sentido, O objetivo do trabalho é simular a iluminação em imagens digitais de média e baixa resolução por meio do desenvolvimento de uma ferramenta. Para se alcançar o objetivo, pretende-se que a ferramenta considere informações de luz em um espaço tridimensional para simular a iluminação na imagem bidimensional. Portanto o que se espera deste trabalho é que imagens de média e baixa resolução apresentem boa iluminação.

Palavras-chave: Iluminação. Algoritmo. Ambiente. Bidimensional.

1 INTRODUÇÃO

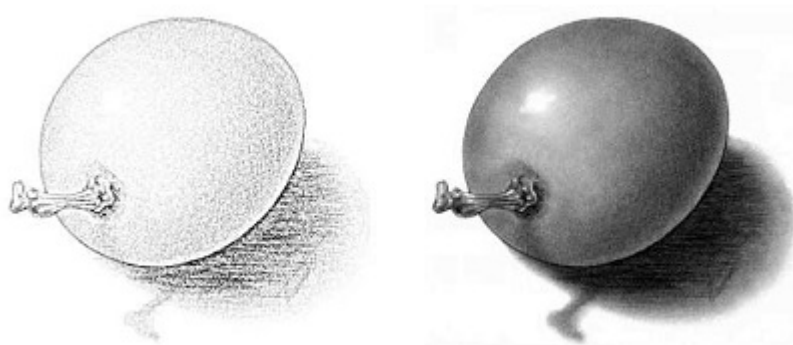
Desde que as pinturas começaram, se formaram vários estilos de arte, ainda que no princípio tudo era voltado a chegar ao realismo das pinturas, por não existir nada que pudesse retratar rostos e momentos melhor do que as artes.

No entanto, isso mudou quando surgiram as câmeras, que poderiam captar tudo de forma praticamente instantânea, com isso era o momento de se pensar que as pinturas iriam sumir, mas ela se renovou e ao invés de buscar o realismo, agora ela estava em busca de trazer novas sensações, como as pinturas abstratas, estilos como cubismo ou surrealismo que usavam de várias curvas e cores para demonstrar expressões (CEZIMBRA, 2021).

Esses estilos artísticos foram uma revolução no mundo da arte, como o surrealismo por exemplo, que tratava de possibilidades infinitas em suas obras, porque ali tudo era possível e o único limite era imaginação, Nesse período foram apresentadas várias obras que ao olhar não faziam sentido algum, porém era esse o princípio da ideia, pois elas eram criadas para passar uma sensação a quem olhasse, como a de liberdade, entre outros.

Atualmente a iluminação é de extrema importância para as pinturas e desenhos digitais pois é ela quem traz o volume ao ambiente, como pode ser observado na Figura 1.

Figura 1 – Duas imagens demonstrando a diferença da iluminação



(GRAFITTI, 2019)

Como podemos observar que na figura acima, a imagem da esquerda apresenta apenas um rascunho, sem uso das técnicas de iluminação a portanto a noção de volume e profundidade não fica muito claro, por outro lado a imagem da direita sendo algo mais agradável nas pinturas como é possível notar, além do mais várias composições de cores da luz retratada em uma imagem pode trazer vislumbres fantásticos.

Com a chegada dos primeiros modelos em três dimensões, como cubos, pirâmides e esferas entre outros, foi possível enxergar através de uma tela, um objeto qualquer que fosse e isso mudou tudo, hoje podemos ver vários exemplos de modelos que está em nossa

realidade, como a impressora 3D ou o *Billboard*¹, pois é com ela que foi possível ver a grande beleza dos modelos, é por esse motivo pelo qual atualmente a iluminação é essência para nos causar a sensação de volume, sem pensar nos algoritmos mais robustos como *ray tracing* que segundo Shirley e Morley (2008) retrata, é um algoritmo onde através de uma janela, os raios são direcionados para as imagens. As superfícies são perdidas ou atingidas por cada pixel, que é representado por um raio. Ao atingir uma superfície, o raio se refrata e continua em um novo curso, causando a formação de luzes adicionais nos arredores.

Porém para os artistas chegarem perto dos algoritmos de iluminação que criam visuais impecáveis em três dimensões, as pinturas e artes foram obrigadas a criar um volume e melhorar a iluminação, pois como a luz funcionava de forma sistemática em três dimensões, as artes puderam abrir espaço em uma variação de cores e paletas diferenciadas que trazem um dinamismo maior.

Por isso atualmente a iluminação se tornou algo tão importante para o mundo, que não importa mas o estilo que é usado, é uma parte essencial. É nesse momento em que chegamos ao cerne do problema, pois com essa importância que a iluminação tem sobre a arte e os avanços tecnológicos cada vez mais ligados a ter um *design* rápido e com eficiência, ferramentas que criam a luz em pinturas, desenhos, objetos, *sprites* ou até logos poderia impactar a arte de maneira eficiente e até mesmo para aumentar a comunidade, fazendo com que várias pessoas iniciantes que não conseguem bons resultados possam aparecer no mundo do *design*.

1.1 Objetivos

1.1.1 Objetivo Geral

O objetivo do trabalho é simular a iluminação em imagens digitais de média e baixa resolução por meio do desenvolvimento de uma ferramenta.

1.1.2 Objetivos Específicos

- Encontrar algoritmos que têm o propósito de iluminar cenas bidimensionais;
- Criar um protótipo e desenvolver a ferramenta para iluminação;
- Realizar testes na ferramenta Relight que será analisado para comparação dos resultados.

¹ apresentado em Tóquio, uma televisão imensa que pode captar imagens como se fossem 3D, criando um aspecto de profundidade pela sua curvatura

2 REVISÃO DA LITERATURA

Este capítulo será organizado de forma a se iniciar com as formas mais primitivas e os modelos de iluminação do mais simples para o mais complexo, logo após mostrar as técnicas de sombreamento e por fim mostrar as ferramentas e trabalhos correlatos ao proposto neste

2.1 Modelos de Iluminação

Segundo Azevedo, Conci e Vasconcelos (2022) o tratamento da luz é um passo essencial para visualizações realistas em ambientes tridimensionais, porém mais do que isso, os processos para chegar a tal são cada vez mais complexos conforme avançam os sistemas de hardware.

Se pensarmos em como funciona a luz no mundo real, é somente através dela que podemos enxergar qualquer coisa, pois mesmo a noite ainda recebemos a luz do sol refletida pela lua, ou seja, tudo o que vemos é somente porque os raios de luz atingem a nossa visão (MANSSOUR; COHEN, 2006).

A princípio, a iluminação de forma geral necessita de uma fonte, algo que possa emitir luz, como por exemplo o sol as estrelas, uma vela ou até mesmo um fósforo. tais fontes criam uma aura ao redor que é conhecida como radiação eletromagnética, e podem emitir entre 380 a 750 nanômetros. Azevedo, Conci e Vasconcelos (2022)

Porém apenas com fontes de luz a cena não se torna interessante, por isso se iniciou a iluminação de forma a ser um ponto tão longe no espaço que ilumina todos os ambientes, e para esse tipo de iluminação surgiram vários modelos.

2.1.1 Reflexão Ambiente

A luz ambiente ou reflexão ambiente é um modelo simplificado para entendermos a luz, ela utiliza padrões da realidade de forma a possibilitar a visualização de cores através da cor do objeto.

Ela "representa um nível constante de luz que define a silhueta dos objetos. Esse nível constante é a simplificação do modelo de iluminação global proposto por Goral et al. (1984) que considera múltiplas reflexões da luz nos diversos objetos que compõe a cena"(SCALCO, 2012, p. 150)

De forma a esclarecer graficamente a reflexão ambiente é possível observar o modelo

sendo renderizado¹ através da seguinte Figura 2.

Figura 2 – Renderização usando apenas a componente ambiente.



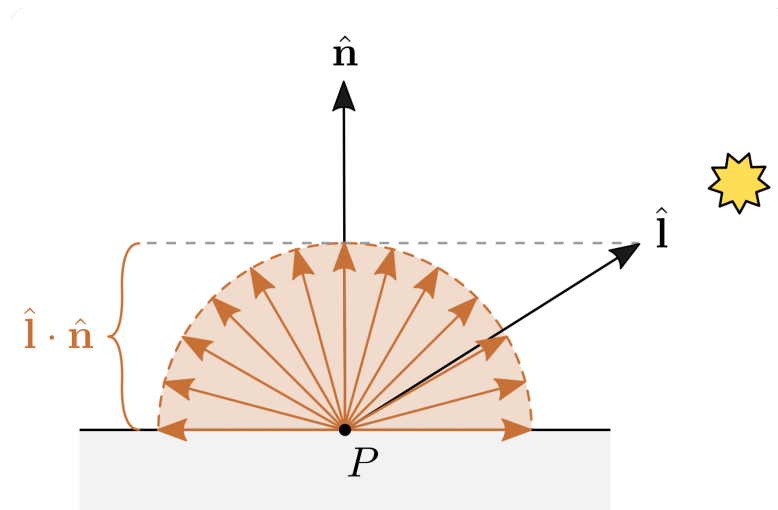
Fonte: (BATAGELO, 2021)

2.1.2 Reflexão Difusa

A reflexão difusa não deixa resultado sombras no ambiente, isso acontece porque a luz difusa provém de diferentes posições criando um volume no objeto sem sombras a sua volta, essa ideia foi feita a partir do Alcalde (2012) e um estudo de luzes em superfícies rugosas.

Matematicamente, falando podemos utilizar dos estudos de (MARSCHNER; SHIRLEY, 2015) Quando uma superfície lambertiana² é iluminada, a luz que incide sobre o ponto P é refletida uniformemente em todas as direções. Isso é ilustrado na figura 3 a seguir.

Figura 3 – Reflexão difusa ideal.



Fonte: (BATAGELO, 2021)

¹ É o processo pelo qual se obtém imagens digitais resultantes de modelos tridimensionais

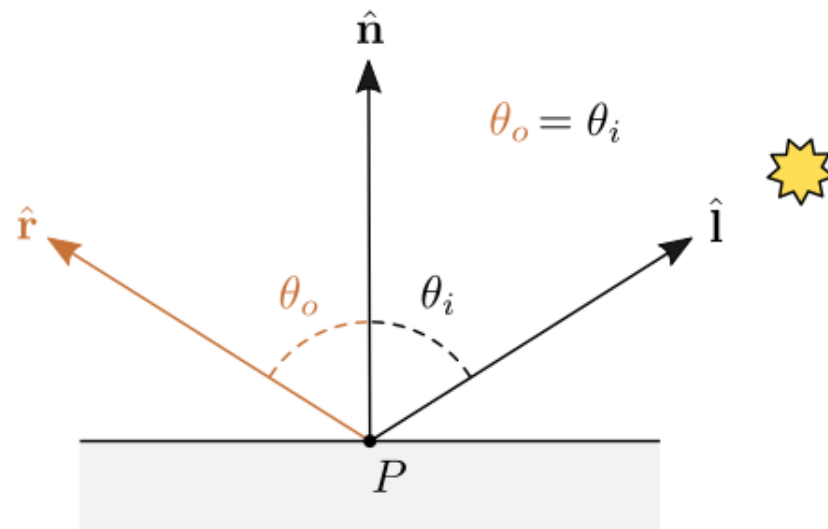
² De forma controlada e uniforme, a superfície lambertiana reflete uniformemente a luz, independentemente da direção em que a luz é refletida.

2.1.3 Reflexão Especular

"a iluminação especular identifica os realces especulares brilhantes que ocorrem quando a luz atinge uma superfície de objeto e reflete de volta em direção à câmera. A iluminação especular é mais intensa do que a luz difusa e incide mais rapidamente na superfície do objeto." dito por (Steven White, John Parente , Michael Satran , Alexander Koren, 2022)

É possível dizer que a partir de uma superfície perfeitamente especular é semelhante a um espelho ideal. Se a luz atinge o ponto P, ela é refletida apenas na direção reflexa \hat{r} do vetor \hat{l} em torno de \hat{n} . Isso é ilustrado na Figura 4.

Figura 4 – Reflexão especular ideal.



Fonte: (BATAGELO, 2021)

2.2 Sombreamento

De acordo com Batagelo (2021) "Sombreamento ou tonalização (do inglês *shading*) é o processo de modificar a intensidade das cores de uma imagem através de tons claros e escuros, de modo a produzir a percepção de volume e profundidade de um objeto tridimensional."

2.2.1 Flat Shading

O sombreamento plano está em sua eficiência. A frequência de cálculo do modelo de iluminação é diretamente proporcional ao número de faces, que geralmente é menor que o número de *pixels*, resultando em menos avaliações.

Segundo o Marschner e Shirley (2015) O sombreamento plano geralmente expõe as arestas irregulares de uma malha poligonal, resultando em uma aparência fragmentada.

Isso pode ser um efeito atraente se o objeto em questão for um poliedro. No entanto, se a malha pretende aproximar uma superfície lisa, o resultado desejado só pode ser alcançado dividindo a malha em faces menores. Neste exemplo, uma malha triangular atua como uma imitação de uma esfera. Ao aumentar o número de subdivisões, é fácil ver que os intervalos de luz entre as faces são equilibrados suavemente.

Atualmente, o uso de sombreamento plano na GPU³ não melhora a eficiência da renderização. Isso ocorre porque o pipeline gráfico atual é otimizado para lidar com atributos de vértice em vez de atributos de triângulo.

2.2.2 Sombreamento de Gourand

A suavização das transições de luz entre as faces dá ao sombreamento Gouraud uma vantagem sobre a técnica de sombreamento plano, pois melhora significativamente o aspecto visual de um objeto, superando visualmente sua contraparte com igual número de faces. Embora digno de nota, o efeito visual apelidado de bandas de Mach⁴ geralmente faz com que as bordas dos rostos pareçam mais pronunciadas do que a área interna. Isso é importante para que as imagens possam ser retratadas como objetos realistas em desenhos, esse efeito resulta do contraste exagerado criado nas bordas, que o sistema visual humano detecta facilmente com base na obra de Gourand (1971)

Onde faltam vértices de malha, pode haver um efeito de brilho cintilante, causado pela perda de brilho especular. Outro problema que pode ser enfrentado com o sombreamento Gouraud é essa ocorrência. Para aliviar isso, pode-se aumentar o número de refinamentos de malha, embora isso possa aumentar o tempo de processamento. Se as subdivisões forem baixas, há uma chance de encontrar esse problema como mostrado por Batagelo (2021)

Calcular o vetor normal do vértice é uma etapa crucial na utilização do sombreamento de Gouraud, pois ele pode ser obtido tomando-se a média dos vetores normais da face próxima.

2.2.3 Sombreamento de Phong

Segundo (??) No método de sombreamento de Phong, a partir das normais aos vértices, realizada do mesmo modo que o sombreamento de Gouraud retratado abaixo, é calculada a normal a cada quadrícula através da interpolação das normais.isto é então usada no modelo de iluminação de Phong para calcular a intensidade da energia luminosa reflectida. Batagelo (2021) Explica que o sombreamento de Phong é que exige mais dos sombreamentos, e consiste em avaliar a equação do modelo de iluminação para cada

³ *Graphics Processing Unit*, ou unidade de processamento de gráficos

⁴ O físico Ernst Mach se deparou com esse fenômeno na década de 1860, que pode ser atribuído a um contraste elevado das bordas em tons de cinza.

fragmento. O resultado é muito superior ao sombreamento de Gouraud para um mesmo número de faces. Marschner e Shirley (2015) retrata que para o modelo de Phong precisamos calcular um vetor \hat{n} para cada fragmento. Isso geralmente ocorre da interpolação linear das coordenadas (x,y,z) dos vetores normais de vértice, igual aos componentes RGB das cores dos vértices são interpoladas para criar um degradê de cores. Porém, neste caso, o vetor de coordenadas interpoladas é normalizado novamente no *fragment shader*⁵ para ser utilizado como vetor \hat{n} da equação.

No sombreamento de Phong, as bandas de Mach são praticamente imperceptíveis. Além disso, o brilho especular é mantido independentemente do refinamento da malha.

2.3 Segment Anything em Visão Computacional

O modelo Segment Anything (SA) é um avanço crucial na segmentação de imagens, capaz de produzir mais de um bilhão de máscaras em 11 milhões de imagens. Sua flexibilidade permite sua aplicação em diversas áreas, como a imagem médica e a glaciologia, devido à sua capacidade de executar tarefas zero-shot, muitas vezes superando métodos supervisionados. O SA utiliza redes neurais convolucionais (CNNs), uma arquitetura codificadora-decodificadora e mecanismos de atenção espacial para alcançar alta eficiência. No entanto, sua complexidade representa um desafio para implementação em tempo real e em ambientes com recursos limitados (KIRILLOV et al., 2023). Em campos específicos, como microscopia e glaciologia, o SA foi adaptado para melhorar a qualidade da segmentação em diversas condições de imagem, evidenciando sua robustez e versatilidade (YUAN, 2024). Por exemplo, sua adaptação para dados de microscopia levou a melhorias significativas na segmentação e rastreamento de objetos em múltiplas dimensões, destacando o potencial do modelo para resolver problemas complexos em análise de imagens (ARCHIT et al., 2023).

2.3.1 Codificador de Imagens (Image Encoder)

O codificador de imagens utilizado pode ser qualquer rede que produza uma incorporação (embedding) da imagem com dimensões $C \times H \times W$. Para garantir escalabilidade e acesso a um pré-treinamento robusto, é empregado um Vision Transformer (ViT) pré-treinado com MAE, especificamente um ViT-H/16. Este ViT-H/16 possui atenção em janelas de 14×14 e quatro blocos de atenção global igualmente espaçados. A saída do codificador é uma incorporação reduzida em $16\times$ da imagem de entrada, a qual possui uma resolução de entrada de 1024×1024 , resultando em uma incorporação final de 64×64 .

Para reduzir a dimensão dos canais, é aplicada uma convolução 1×1 para alcançar 256 canais, seguida de uma convolução 3×3 também com 256 canais. Ambas as convoluções

⁵ através da openGL

são seguidas por uma normalização em camadas (*Layer Normalization*).

2.3.2 Codificador de Prompts (Prompt Encoder)

Prompts esparsos são mapeados para incorporações vetoriais de 256 dimensões. Pontos são representados como a soma de uma codificação posicional da localização do ponto e de uma incorporação aprendida que indica se o ponto está no primeiro plano ou no fundo. Caixas são representadas por um par de incorporações: uma para o canto superior esquerdo e outra para o canto inferior direito, utilizando codificações posicionais e incorporações aprendidas.

Para representar texto livre, é utilizado o codificador de texto do CLIP, com a possibilidade de utilizar qualquer codificador de texto. Prompts densos, como máscaras, possuem correspondência espacial com a imagem e são inseridos em uma resolução $4\times$ menor do que a imagem de entrada. Estas máscaras são então reduzidas adicionalmente $4\times$ usando convoluções 2×2 , resultando em uma dimensão de canal final de 256, com ativações GELU e normalização em camadas entre as etapas.

2.3.3 Decodificador de Máscaras (Lightweight Mask Decoder)

O decodificador de máscaras é um módulo eficiente que mapeia a incorporação da imagem e um conjunto de incorporações de prompt para uma máscara de saída. Para combinar essas entradas, nos inspiramos em modelos de segmentação baseados em Transformer (CARION et al., 2020)(CHENG; SCHWING; KIRILLOV, 2021) e modificamos um decodificador Transformer padrão (VASWANI et al., 2017). Antes de aplicar nosso decodificador, inserimos no conjunto de incorporações de prompt um token de saída aprendido, que será utilizado na saída do decodificador, de forma análoga ao token [class] descrito em (VASWANI et al., 2017). Para simplificar, referimos a essas incorporações (excluindo a incorporação da imagem) coletivamente como “tokens”.

Para garantir que o decodificador tenha acesso a informações geométricas críticas, as codificações posicionais são adicionadas à incorporação da imagem sempre que participam de uma camada de atenção. Além disso, os tokens de prompt originais completos (incluindo suas codificações posicionais) são re-adicionados aos tokens atualizados sempre que participam de uma camada de atenção. Isso permite uma forte dependência tanto da localização geométrica quanto do tipo do token de prompt.

Após a execução do decodificador, ampliamos a incorporação da imagem atualizada em $4\times$ usando duas camadas convolucionais transpostas (resultando em uma escala $4\times$ menor em relação à imagem de entrada). Em seguida, os tokens realizam uma última atenção para a incorporação da imagem, e passamos a incorporação do token de saída atualizada para uma pequena MLP de três camadas que gera um vetor correspondente

à dimensão de canal da incorporação da imagem ampliada. Finalmente, prevemos uma máscara através de um produto ponto a ponto espacial entre a incorporação da imagem ampliada e a saída da MLP.

O Transformer utiliza uma dimensão de incorporação de 256. Os blocos MLP do Transformer têm uma grande dimensão interna de 2048, mas a MLP é aplicada apenas aos tokens de prompt, que são relativamente poucos (raramente ultrapassando 20). No entanto, em camadas de cross-atenção, onde temos uma incorporação de imagem de 64×64 , reduzimos a dimensão do canal das queries, keys, e values em $2 \times$ para 128, visando eficiência computacional. Todas as camadas de atenção utilizam 8 cabeças.

As convoluções transpostas usadas para ampliar a incorporação da imagem de saída são 2×2 , com stride 2 e dimensões de canal de saída de 64 e 32, e utilizam ativações GELU. Elas são separadas por normalização em camadas.

2.4 Geração de Mapas de Normais

Os mapas de normais são essenciais para definir como a luz interage com as superfícies, criando efeitos de iluminação dinâmica em gráficos 3D, pixel art e renderização de fluidos. A geração de mapas de normais, especialmente para pixel art, é desafiadora devido à sua estética única. Métodos como amostragem geométrica e modelos generativos profundos, como Redes Adversárias Generativas Condicionais (CGANs), foram desenvolvidos para gerar mapas de normais precisos. Um estudo compilou diferentes técnicas de geração de mapas de normais e analisou sua aplicabilidade ao pixel art, contribuindo para uma análise qualitativa do comportamento dessas técnicas em diferentes cenários (MOREIRA FLÁVIO COUTINHO, 2013). No contexto de renderização de fluidos, a refinada representação dos mapas de normais por meio de redes neurais aprimora a qualidade da renderização, mantendo detalhes e suavidade. O uso de CGANs tem demonstrado eficácia em melhorar a precisão dos mapas de normais, capturando detalhes intrincados e aprimorando a iluminação realista em renderizações (CHOI et al., 2021). A integração dessas técnicas na ferramenta proposta no seu TCC pode aprimorar a percepção de profundidade e a iluminação realista em aplicações de imagens.

2.4.1 Métodos Tradicionais

Os mapas de normais são amplamente utilizados para simular detalhes de superfícies sem a necessidade de adicionar geometria adicional a um modelo 3D. Um método tradicional de gerar mapas de normais é a partir de malhas 3D de alta resolução. Nesta abordagem, um modelo detalhado é projetado em uma malha de menor resolução, e a diferença nas inclinações das superfícies é codificada em um mapa de normais. Outro método convencional é o bump mapping, que utiliza mapas de normais para simular rugosidades

e imperfeições na superfície de um objeto 3D, proporcionando a ilusão de detalhes sem alterar a malha subjacente. Embora essas técnicas sejam eficazes, elas possuem limitações, especialmente quando aplicadas a estilos gráficos mais estilizados, como o pixel art, ou quando a resolução das imagens é baixa. Em tais casos, os métodos tradicionais podem não capturar corretamente as nuances da superfície ou podem introduzir artefatos visuais, resultando em uma renderização menos realista. Por exemplo, a análise de (MOREIRA FLÁVIO COUTINHO, 2013) destaca que a aplicação desses métodos em pixel art pode não produzir resultados precisos devido às especificidades do estilo. Essas desvantagens motivaram o desenvolvimento de técnicas mais avançadas, que buscam superar os limites das abordagens tradicionais e oferecer soluções mais adaptáveis a diferentes estilos e cenários.

2.5 Técnicas avançadas de pixel arte

A geração de mapas de normais para pixel art apresenta desafios únicos devido à natureza estilizada e à baixa resolução desse tipo de arte. Métodos tradicionais muitas vezes falham em capturar os detalhes necessários para representar corretamente a interação da luz com a superfície pixelada. Uma técnica avançada para abordar esse problema é a amostragem geométrica, que se concentra na análise da forma geométrica dos pixels para criar mapas de normais mais precisos 2021. Além disso, o uso de Redes Adversárias Generativas Condicionais (CGANs) oferece uma abordagem inovadora para a geração de mapas de normais em pixel arte, permitindo a produção de resultados mais refinados e detalhados. Por exemplo, 2018 desenvolveram um sistema interativo que utiliza técnicas de deep learning para gerar mapas de normais de alta qualidade a partir de esboços, melhorando a precisão e a interatividade do processo. Essas redes utilizam amostras geométricas para treinar modelos capazes de gerar mapas de normais que respeitam as características visuais do pixel arte, como bordas nítidas e transições suaves de luz. No contexto de renderização de fluidos, abordagens baseadas em aprendizado profundo, como a proposta por 2021, refinam os mapas de normais para aprimorar a qualidade da renderização, preservando detalhes e suavidade. Essas técnicas avançadas são essenciais para garantir que a iluminação e a profundidade sejam representadas de forma precisa, resultando em uma renderização mais realista e visualmente atraente.

2.6 Ferramentas para iluminação de cenários

Para a iluminação ser feita em ambiente gráfico foram criadas diversas ferramentas, dentre elas Vulkan e OpenGL, que trabalham em cima de toda a forma como a luz e os materiais funcionam no ambiente, proporcionando configurações avançadas de baixo nível para que exista controle total do ambiente sendo desenvolvido.

2.6.1 Vulkan

Segundo o (Khronos Group, 2016), a Vulkan⁶ é uma API ⁷ de gráficos de baixo nível, criada pela Khronos Group, com o objetivo de fornecer aos desenvolvedores um acesso mais direto ao hardware do computador para criar aplicativos gráficos de alto desempenho em várias plataformas e dispositivos. Ela pode ser usada para desenvolver aplicativos para diversos casos de uso. Embora as aplicações em Vulkan possam escolher usar um subconjunto das funcionalidades descritas abaixo, a API foi projetada para que um desenvolvedor possa usar todas elas em uma única API.multiplataforma.

2.6.2 OpenGL

OpenGL⁸ é uma biblioteca de renderização. No entanto, o que o OpenGL não faz é reter informações sobre um "objeto". Tudo o que o OpenGL vê é uma esfera de polígonos e um conjunto de estados com os quais renderizá-los. não se lembra que foi desenhado uma linha em um local determinado e uma esfera em outro.

Por causa disso, a maneira geral de usar o OpenGL é desenhar tudo o que você precisa desenhar e, em seguida, mostrar essa imagem com um comando de troca de *buffer* dependente da plataforma. Se você precisar atualizar a imagem, desenhe tudo novamente, mesmo que precise atualizar apenas parte dela. Se você quiser animar objetos se movendo na tela, precisa de um *loop* que constantemente limpe e redesenhe a tela.

2.6.3 Engines

Está sessão tem o intuito de apresentar as *engines* do mercado atual, afinal elas tem sua própria forma de criar iluminação e alterar de diversas maneiras para atender a um estilo visual único com *shaders* ou outros efeitos.

De acordo com (FELIPE, 2017) Em sua essência, a *engine* é um *software* vital que compila todos os arquivos e bibliotecas essenciais dos quais um jogo depende. Essa ferramenta serve como base para o desenvolvimento de jogos, permitindo que os desenvolvedores criem habilmente sua visão. No entanto, a criação de jogos requer muito mais do que apenas um mecanismo. Os desenvolvedores também devem fazer uso de editores de imagem, software de áudio e vídeo, modeladores 3D e software vetorial, dependendo da situação. Uma vez criados, esses elementos ganham vida dentro do motor, onde recebem animação, física, efeitos sonoros e outros recursos cruciais.

⁶ disponível: <https://www.vulkan.org/>

⁷ Application Programming Interface

⁸ disponível: <https://www.opengl.org/>

2.6.3.1 Unity

A marca registrada do *Unity* reside em seu estilo de programação, utilizando de vários objetos conectando vários pontos em um mesmo código. Distinto e estrutura de projeto que possui uma simplicidade inigualável. O *Unity* permite que os desenvolvedores utilizem as alternativas disponíveis e abre vários caminhos para os criadores se concentrarem em seus conhecimentos, principalmente na conduta dos personagens, de ambas as classes segundo (Henrique, 2014)

Embora o *Unity* tenha um objetivo de desenvolvimento bem definido, seus recursos se estendem além disso para diversos tipos de projeto. Gráficos altamente realistas são a combinação perfeita para jogos de aventura como RPGs ⁹, TPSs ¹⁰ e FPSs ¹¹ ao usar o *Unity*. A capacidade de incorporar elementos feitos por outros desenvolvedores de jogos aos nossos é um dos maiores pontos fortes do *Unity*. Essa funcionalidade é tremendamente vantajosa para indivíduos sem habilidades gráficas extensas, como modelagem 3D ou ilustração.

2.6.3.2 Godot

De acordo com Juan Linietsky, Ariel Manzur and the Godot community (2014), na Godot você tem liberdade criativa para personalizar o código-fonte do mecanismo de jogo, totalmente adaptado às suas necessidades. O software é flexível no sentido de que pode facilmente gerar jogos 2D e 3D, com uma abundância de recursos como gráficos, som, física e animação. Além disso, sua compatibilidade entre plataformas se estende a *Windows*, *macOS*, *Linux*, *Android*, *iOS* e *web*.

Comparado a outros mecanismos de jogos de código aberto, o Godot se destaca por sua interface amigável e processo de aprendizado simples. Com documentação abrangente, os aspirantes a criadores de jogos podem mergulhar direto no processo de criação do jogo. Os jogadores também podem fazer sua programação em várias linguagens diferentes, incluindo CSharp e GDScript. Esse recurso torna mais fácil para os desenvolvedores selecionarem o idioma de sua preferência.

2.7 Trabalhos Correlatos

Nesta seção são apresentados os trabalhos correlatos ao proposto neste trabalho

⁹ Role Playing Game ou Jogo de interpretação de personagem

¹⁰ Third Person Shooter ou Tiro em terceira pessoa

¹¹ First Person Shooter ou Tiro em primeira pessoa

2.7.1 Geração De Mapas De Iluminação Baseado Em Topologia Estimada Para Iluminação De Sprites

Este trabalho de autoria Santana (2017) trata-se de uma pesquisa prática em torno do mapeamento de uma imagem, e criação de uma topologia, no âmbito das pixel artes atualmente esse processo de iluminação é realizada de duas formas, com modelos em 3D ou de forma manual, mas o que realmente é tratado desse trabalho é de como esse tipo de tecnologia pode facilitar a vida do profissional.

Em seu desenvolvimento é retratado métodos já criados para tentar resolver o problema, um deles é o *Cross Shade* proposto por (SHAO et al., 2012) que cria em um desenho uma especie de cruz para que em cada plano do desenho possa ser identificado a sua normal de forma como a cruz está, pegando sua curvatura para identificar um objeto em três dimensões

Dentre os métodos é utilizado para os resultados uma criação de mapa topológico a partir da ferramenta *sprite illuminator* e algumas customizações por cima do que foi gerado para que fosse possível uma bom resultado em pixel artes bem pequenas.

2.7.2 Analysis and Compilation of Normal Map Generation Techniques for Pixel Art

Esse trabalho do autor Moreira Flávio Coutinho (2013) tem como influencia a luz que interage com cada pixel do material. Existem vários métodos para gerar mapas normais em jogos 3D, mas aplicá-los na arte pixel pode não produzir resultados precisos por causa das particularidades do estilo. Este trabalho reúne diferentes métodos de geração de mapas normais e estuda sua aplicação na arte pixel, contribuindo para uma análise qualitativa do comportamento desses métodos em diferentes estudos de caso e diminuindo a escassez de material existente sobre essas técnicas.

2.7.3 CrossSketch: freeform surface modeling with details

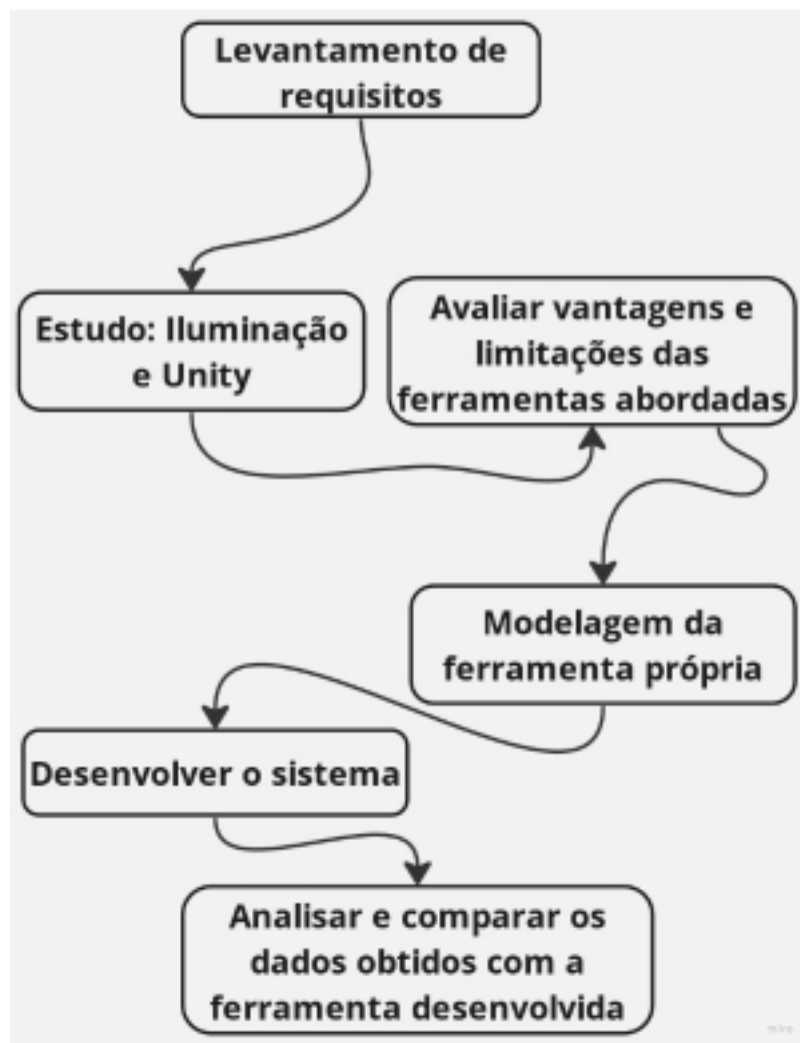
Este artigo do autor Andre e Nakajima (2007) apresenta uma nova técnica para modelar uma superfície tridimensional livre, incluindo sua forma global e pequenos detalhes, usando uma interface de desenho a partir de um único ponto de vista. Em sistemas de modelagem anteriores que utilizavam esboços como entrada, a forma era reconstruída a partir da silhueta e o usuário tinha pouco controle sobre as partes internas do resultado. É gerado uma grade de linhas co-planas a partir de um pequeno número de traços desenhados pelo usuário, foi estimado o vetor normal onde ele é restrito e formamos a superfície propagando essa informação para toda a grade. Como resultado, traços menores atuam localmente para adicionar detalhes, enquanto traços maiores modificam toda a superfície. Esse trabalho oferece uma nova abordagem para o problema de modelagem

a partir de esboços e tem a intenção de ser parte de um sistema de modelagem mais complexo.

3 METODOLOGIA

Descrever metodologia, materiais e métodos utilizados no estudo, bem como os procedimentos experimentais realizados, nesta etapa será descrito vários assuntos sobre os passos a serem realizados iniciando com o levantamento dos dados, para uma análise e estudo posteriormente analisando todos os processos que são realizados, afim de iniciar o desenvolvimento.

Figura 5 – Passos feito na plataforma miro



Fonte: Feito pelo autor

3.1 Levantamento de Requisitos

Nessa sessão será abordado a pesquisa qualitativa feita a partir da ferramenta *Relight* que usa de um algoritmo para identificação de camadas em uma imagem, assim,

quando a imagem é colocada na ferramenta ela faz um mapeamento das áreas altas e baixas para criar as camadas, por esse motivo as camadas que são denominadas como mais alta recebe mais luz do que as mais baixas.

Porém a ferramenta tem uma funcionalidade a mais onde é possível definir em que nível está a luz, fazendo com que a iluminação do objeto possa vir das camadas mais a baixo para as camadas acima, está ferramenta com a luz funciona como um pincel de adicionar em softwares de pintura, pois ela recebe a luz já existente na foto e complementa por cima com a luz própria, podendo ser de qualquer cor.

A seguir será mostrado na tabela, onde terá a avaliação qualitativa referente a vários tipos de imagens testadas a partir da ferramenta, que para efeitos deste trabalho serão utilizados os valores de alta resolução como acima de 720 pixels, média resolução entre 720 e 256 pixels e baixa resolução abaixo de 256 pixels, retratado na tabela:

Avaliação qualitativa da ferramenta

Imagem	Alta Resolução	Média Resolução	Baixa Resolução
Objetos	Funciona	Funciona	Restrição
Humanos	Funciona	Funciona	Restrição
Paisagens	Funciona	Problema	Problema

Nessa tabela, foi adotado três categorias para descrever o funcionamento: Funciona, Restrição e Problema. A primeira categoria é utilizada para descrever que a imagem utiliza se comportou de forma correta, Restrição é a categorização para as imagens que funcionam mas em alguns casos ou áreas da imagem geram problemas, já a categoria Problema indica os casos onde a luz não reconhecia as formas de maneira correta.

3.2 Estudo: Iluminação e Unity

Nesse estudo será demonstrado algumas das partes que serão necessárias para o desenvolvimento da ferramenta, como a seguir que se iniciará com a iluminação nas histórias em quadrinhos. Sempre que vemos uma cor clara e logo depois uma escura, isso significa que há um relevo muito grande ali, por exemplo na Figura 6

Nesse quadrinho do Batman, podemos ver que em seu ombro está muito claro e logo acima onde sua capa está vem uma cor muito escura, isso nos mostra que a profundidade da capa é grande, ao ponto de não chegar nenhuma luz até ela, é claro que precisamos levar em consideração que nas HQ's em geral os contrastes são muito maiores, porque traz esse volume nos trajes.

Partindo para o ramo da *Unity* será preciso um aprendizado todo relacionado a iluminação dentro da ferramenta, uso dos objetos em cena e manipulação deles, pois para o desenvolvimento será necessário a criação de uma malha que forme a imagem, um

Figura 6 – História em quadrinhos, Batman



Fonte: (BAZELA, 2022)

sistema de camadas para que possa estabelecer profundidade, um sistema de alteração e manipulação da posição, cor e luminosidade do ponto de luz. E por fim um estudo básico de toda interface da *Unity*.

Outro ponto importante no estudo, será a linguagem de programação **C#** que é usada como alicerce para qualquer código que precise ser estruturado lá dentro, desde instanciar objetos dentro da cena até modificar configurações de câmera como movimento, posicionamento e ângulo até para recebimento dos arquivos

3.3 Avaliar vantagens e limitações das ferramentas abordadas

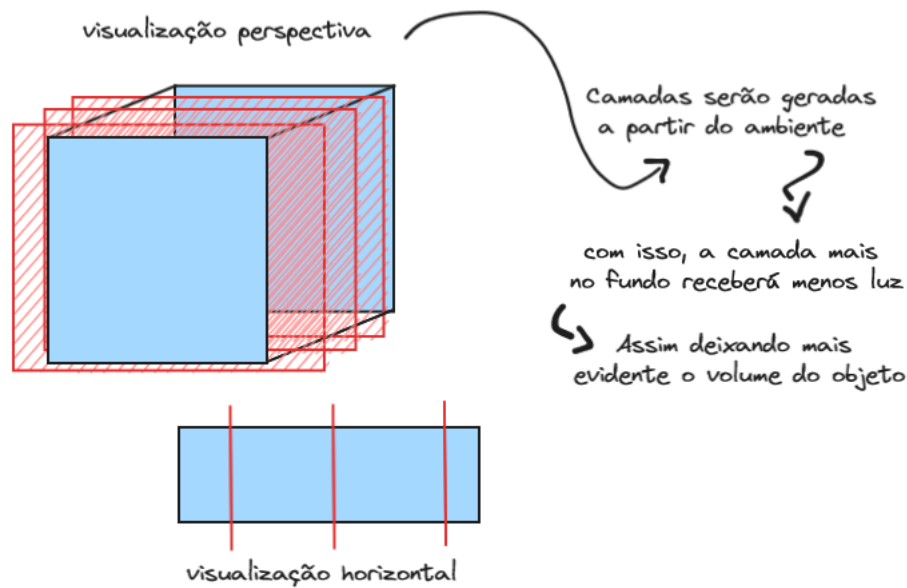
O *Relight* é uma ferramenta muito boa para processos de iluminação tanto para cenários como personagens, com uma boa modificação como profundidade, posição, cor e luminosidade, dando para o usuário a liberdade de dar personalidade as suas imagens.

Mas ainda existe várias limitações e elas que serão retratadas nessa seção, como foi visto na análise a cima na tabela é possível notar que quanto menor a imagem é, menos funcional ela passa a ser, por exemplo imagens em baixa resolução ou pixeladas começam a ter muitos problemas pois a luz não consegue distinguir os objetos na cena e muito menos paisagens pois com tantas cores perto uma das outras transforma a cena em uma desordem visual.

3.4 Modelagem da ferramenta

Uma malha deverá ser criada por cima de toda a imagem e definir um inteiro para a profundidade, cada *pixel* da imagem deve receber uma profundidade como pode se observar na Figura 7

Figura 7 – Rascunho elaborado pelo autor



Assim ele vai numerar toda a malha de *pixels* na tela com números que podem ter uma variação maior dependendo do tamanho da imagem (no caso de exemplo)

Será calculado de 1 a 9, mas com imagens imensas em alta resolução podemos pensar em usar de 1 a 100 ou até mais

Podemos pensar que dessa forma será possível utilizar essa profundidade para produzir uma luz pelas laterais ou pela frente e até atrás de elementos, sem perder sua funcionalidade até mesmo em imagens pequenas

Além disso, cada um desses valores definidos em cada *pixel* da imagem, vão servir de referencia para criação de vários objetos, ele levará em consideração a distância e a variação de cor que foi atingida

Como por exemplo, se existir um objeto a frente da camada 1 até a 10 e um objeto atrás que está na camada 15, essa distancia de 5 camadas irá fazer o objeto dá frente se separar com o de trás transformando a cena com dois objetos ao invés de um

3.5 Justificativa das Tecnologias a serem adotadas

Será utilizado a IDE Visual Studio para a criação dos códigos que serão importados na *Unity*, foi optado ele pois é o sistema mais robusto para a utilização do **C#**.

Para o processo de desenvolvimento efetivo, além da IDE será usado a *game engine*, *Unity*, que foi escolhida pela simplicidade em aplicar luz e criação de objetos bidimensionais em ambientes Tridimensionais, ou como é conhecido 2.5D, para assim poder criar com efetividade a malha e a luz aplicada a ela

3.6 Avaliação Qualitativa

Logo após todo o processo estiver concluído, chega a etapa da avaliação, nesse tópico irá apresentar como será feita a avaliação dos resultados obtidos a partir das análises feitas anteriormente, nesta comparação será possível avaliar os seguintes pontos, primeiramente a alteração em imagens de baixa resolução se é possível a partir desse novo algoritmo e se as vantagens irão se permanecer. Outro ponto que poderá ser analisado será a eficácia de executar a ferramenta e se ela é possível a partir dos cálculos necessários.

4 ANÁLISE DOS RESULTADOS PRELIMINARES

Inicialmente como resultado foi obtido a inicial análise feita na plataforma de iluminação *Relight* que nos traz diversos pontos importantes para a criação da ferramenta.

Foi diagnosticado diversas funções e possíveis formas de como elas podem ter sido elaboradas. como por exemplo a utilização de camadas da imagem e a somatório de cores para criar a intensidade de luz sobre o que é apresentado.

4.1 Cronograma do Trabalho

Segue abaixo o cronograma de trabalho das atividades realizadas e das que serão executadas até a Avaliação Final de TCC.

1. Elaboração da Introdução
2. Elaboração da Revisão da Literatura
3. Elaboração da Metodologia
4. Estudo sobre a *Unity* e Iluminação
5. Desenvolvimento da ferramenta
6. Elaboração dos Resultados e Conclusões

[illegible]

5 CONCLUSÕES PARCIAIS

Foi descrita a modelagem da ferramenta própria, que envolve a criação de uma malha por cima da imagem para definir a profundidade e a variação de cor, permitindo a manipulação da luz. Além disso, foi citado as tecnologias adotadas, como o Visual Studio para programação em *C#* e a Unity como a game *engine* escolhida.

Para avaliar a eficácia da ferramenta, foi planejado realizar uma avaliação qualitativa dos resultados obtidos, incluindo a análise da alteração de imagens, a viabilidade da ferramenta e a eficácia dos cálculos realizados. Para fim é esperado que a ferramenta esteja concluída mesmo que de forma manual, para que ocasionalmente possa ter um progresso posterior em outros trabalhos.

REFERÊNCIAS

- ALCALDE, A. K. G. Interacción de luz difusa con superficies rugosas. 2012. 8
- ANDRE, S. S. A.; NAKAJIMA, M. Crosssketch: freeform surface modeling with details. p. 52, 2007. Disponível em: <https://dl.acm.org/doi/abs/10.1145/1384429.1384442?casa_token=ynVUdINw8eEAAAAA:mu8ZoSKq_D8QfR4RwMT4RHnyB6rCD1ScV4kjGh33Ie6rt-kafs7Uked4g0i_zABu9NkJc7RStqhfig>. 17
- ARCHIT, A. et al. Segment anything for microscopy. **bioRxiv**, 2023. 11
- AZEVEDO, E.; CONCI, A.; VASCONCELOS, C. **Computação Gráfica Teoria e Prática: Geração de imagens**. 1. ed. [S.l.: s.n.], 2022. v. 1. 7
- BATAGELO, B. M. H. **Reflexão Especular**. 2021. <<https://www.brunodorta.com.br/cg/lighting.html>>. Accessed: 2023-4-24. 8, 9, 10
- BAZELA, C. **Batman: 3 quadrinhos que influenciaram o novo filme**. 2022. <<https://www.boletimnerd.com.br/batman-3-quadrinhos-que-influenciaram-o-novo-filme/>>. Accessed: 2023-5-20. 21
- CARION, N. et al. End-to-end object detection with transformers. **CoRR**, abs/2005.12872, 2020. Disponível em: <<https://arxiv.org/abs/2005.12872>>. 12
- CEZIMBRA, D. J. **História da Arte e Design**. 1. ed. [S.l.]: Contentus, 2021. 142 p. 5
- CHENG, B.; SCHWING, A.; KIRILLOV, A. Per-pixel classification is not all you need for semantic segmentation. In: RANZATO, M. et al. (Ed.). **Advances in Neural Information Processing Systems**. Curran Associates, Inc., 2021. v. 34, p. 17864–17875. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/2021/file/950a4152c2b4aa3ad78bdd6b366cc179-Paper.pdf>. 12
- CHOI, M. et al. Deep representation of a normal map for screen-space fluid rendering. **Applied Sciences**, Multidisciplinary Digital Publishing Institute, v. 11, n. 19, p. 9065–, 2021. 13, 14
- FELIPE, L. **O que é uma engine? Qual sua importância no desenvolvimento de games?** 2017. Accessed: 2023-4-30. 15
- GORAL, C. M. et al. Modeling the interaction of light between diffuse surfaces. **SIGGRAPH Comput. Graph.**, Association for Computing Machinery, New York, NY, USA, v. 18, n. 3, p. 213–222, jan 1984. ISSN 0097-8930. Disponível em: <<https://doi.org/10.1145/964965.808601>>. 7
- GOURAND, H. **Computer Display of Curved Surfaces**. Tese (Doutorado) — University of Utah, 1971. 10
- GRAFITTI. **Aprenda a criar efeitos de luz e sombra para as suas ilustrações**. 2019. Disponível em: <<https://blog.grafittiartes.com.br/>>

aprenda-criar-efeitos-de-luz-e-sombra-para-as-suas-ilustracoes>. Acesso em: 27 de mar. de 2023. 5

HE, Y. et al. **Sketch-based Normal Map Generation with Geometric Sampling**. 2021. 14

Henrique. **Unity 3D: Introdução ao desenvolvimento de games**. 2014. <<https://www.devmedia.com.br/unity-3d-introducao-ao-desenvolvimento-de-games/30653>>. Accessed: 2023-4-27. 16

Juan Linietsky, Ariel Manzur and the Godot community. **Introdução ao editor do Godot**¶. 2014. <https://docs.godotengine.org/pt_BR/stable/getting_started/introduction/introduction_to_godot.html>. Accessed: 2023-4-27. 16

Khronos Group. **Vulkan Guide**. 2016. <<https://www.vulkan.org/>>. Accessed: 2023-4-30. 15

KIRILLOV, A. et al. Segment anything. **arXiv:2304.02643**, 2023. 11

MANSSOUR, I. H.; COHEN, M. Introdução à computação gráfica. **RITA**, v. 13, n. 2, p. 43–68, 2006. 7

MARSCHNER, S.; SHIRLEY, P. **Fundamentals of Computer Graphics**. [S.l.]: CRC Press, 2015. v. 1. Drawing on an impressive roster of experts in the field, Fundamentals of Computer Graphics, Fourth Edition offers an ideal resource for computer course curricula as well as a user-friendly personal or professional reference. Focusing on geometric intuition, the book gives the necessary information for understanding how images get onto the screen by using the complementary approaches of ray tracing and rasterization. It covers topics common to an introductory course, such as sampling theory, texture mapping, spatial data structure, and splines. It also includes a number of contributed chapters from authors known for their expertise and clear way of explaining concepts. Highlights of the Fourth Edition Include: Updated coverage of existing topics Major updates and improvements to several chapters, including texture mapping, graphics hardware, signal processing, and data structures A text now printed entirely in four-color to enhance illustrative figures of concepts The fourth edition of Fundamentals of Computer Graphics continues to provide an outstanding and comprehensive introduction to basic computer graphic technology and theory. It retains an informal and intuitive style while improving precision, consistency, and completeness of material, allowing aspiring and experienced graphics programmers to better understand and apply foundational principles to the development of efficient code in creating film, game, or web designs. ISBN 978-14-8222-941-7. 8, 9, 11

MOREIRA FLÁVIO COUTINHO, L. C. R. D. Analysis and compilation of normal map generation techniques for pixel art. 2013. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/9961116>>. 13, 14, 17

SANTANA, F. F. G. A. C. D. Geração de mapas de iluminação baseado em topologia estimada para iluminação de sprites. 2017. Disponível em: <<https://app.uff.br/riuff/handle/1/5683>>. 17

SCALCO, R. Uso do geogebra 3d como estúdio de tonalização para análise das componentes dos coeficientes de reflexão do modelo de iluminação local. **Revista do**

Instituto GeoGebra Internacional de São Paulo, v. 1, n. 1, p. XVI – XXVII, mar. 2012. Disponível em: <<https://revistas.pucsp.br/index.php/IGISP/article/view/8198>>. 7

SHAO, C. et al. Crossshade: Shading concept sketches using cross-section curves. **ACM Trans. Graph.**, Association for Computing Machinery, New York, NY, USA, v. 31, n. 4, jul 2012. ISSN 0730-0301. Disponível em: <<https://doi.org/10.1145/2185520.2185541>>. 17

SHIRLEY, P.; MORLEY, R. **Realistic Ray Tracing, Second Edition**. Taylor & Francis, 2008. ISBN 9781568814612. Disponível em: <<https://books.google.com.br/books?id=knpN6mnhJ8QC>>. 6

Steven White, John Parente , Michael Satran , Alexander Koren. **Microsoft Learn: Iluminação Especular**. 2022. <<https://learn.microsoft.com/pt-br/windows/uwp/graphics-concepts/specular-lighting>>. Accessed: 2023-4-24. 9

SU, W. et al. Interactive sketch-based normal map generation with deep neural networks. **Interactive 3D Graphics and Games**, May 2018. Disponível em: <<https://typeset.io/papers/interactive-sketch-based-normal-map-generation-with-deep-1vewh42kks>>. 14

VASWANI, A. et al. Attention is all you need. **CoRR**, abs/1706.03762, 2017. Disponível em: <<http://arxiv.org/abs/1706.03762>>. 12

YUAN, Z. Principles, applications, and advancements of the segment anything model. **Applied and Computational Engineering**, 2024. 11