

## **CURSO DE SISTEMAS DE INFORMAÇÃO**

**ERICK ANDRADE DOS SANTOS**

### **GERAÇÃO DE IMAGENS 3D A PARTIR DO USO DE DUAS WEBCAMS**

ERICK ANDRADE DOS SANTOS RA:0709111204

## GERAÇÃO DE IMAGENS 3D A PARTIR DO USO DE DUAS WEBCAMS

Trabalho de conclusão de curso apresentado à Disciplina Trabalho de conclusão de curso III do Curso de Sistemas de informação da Faculdade de Presidente Prudente – FAPEPE, como requisito principal para avaliação final.

Professor: Danilo Filitto

## DEDICATÓRIA

Dedico este trabalho a minha família por estar presente em todos os momentos e apoiar-me ao longo destes quatro anos.

## AGRADECIMENTOS

Em primeiro lugar eu agradeço a DEUS pelo dom da vida, da sabedoria, da perseverança e do amor, dons estes que alicerçaram todas as minhas realizações.

Agradeço ainda:

A minha família e todos aqueles que contribuíram para minha formação. Em especial: minha mãe Sandra, minha irmã Helen e meu pai Antônio; a professora Erica Cristina (CIEE) e aos professores Valter Bercelli, Danilo Filitto, Paula Langhi, Fernanda Barcaro e todos os outros professores que colaboraram na criação do conhecimento que tenho hoje.

Para que o mau triunfe, basta que os bons não façam nada.  
(Edmund Burke)

## LISTA DE FIGURAS

Figura 1 - Coordenadas $f(x,y)$ .....	14
Figura 2 - Matriz $f(x,y)$ .....	15
Figura 3 - Lena (monocromática).....	15
Figura 4 - Lena (RGB) .....	16
Figura 5 - Pintura de Giotto do século 13.....	18
Figura 6 - Triângulo impossível .....	21
Figura 7 – Paralaxe .....	23
Figura 8 - Tipos Paralaxe.....	24
Figura 9 - Estereoscopia ativa .....	25
Figura 10 - Estereoscopia polarizada.....	26
Figura 11 - Estereoscopia anáglifo .....	27
Figura 12 - Óculos Anáglifo.....	28
Figura 13 - Visão simplificada da classe Câmera (câmera1 id = 0).....	32
Figura 14 - Visão simplificada da classe Câmera 3D.....	34
Figura 15 - Visor da Câmera 3D .....	35
Figura 16 - Controles em estado inicial.....	35
Figura 17 - Controles em estado de transmissão .....	35
Figura 18 - Visão geral simplificada .....	36
Figura 19 - Materiais utilizados no experimento.....	37
Figura 20 - Câmeras já ajustadas.....	37
Figura 21 - Livro distancia 30cm .....	38
Figura 22 - Livro distancia 103cm .....	38
Figura 23 - Livro aberto distancia 74cm .....	39
Figura 24 - CD com reflexo da luz 24cm .....	39
Figura 25 - Copo de vidro 25cm.....	40
Figura 26 - Embalagem 24cm .....	40

## LISTA DE TABELAS

Tabela 1 - Comprimento de onda (Faixa visível) .....	13
Tabela 2 - Classes Emgucv utilizadas .....	30
Tabela 3 - Estrutura de dados simplificada .....	30

## **RESUMO**

O objetivo desse estudo é fornecer uma visão introdutória para área de tecnologias 3D, de modo que possa servir como base de estudo para iniciantes na área ou como referencia para estudos mais avançados. O estudo está dividido em duas partes: uma parte principal contemplando os fundamentos teóricos e uma parte complementar descrevendo a aplicação construída durante o estudo. A parte de fundamentos apresenta o processo de formação de imagem digital, bem como uma breve introdução ao processamento de imagens e tipos de tecnologias 3D, juntamente com os fatos que marcaram a história da mesma. Na parte complementar é descrito de forma não técnica a estrutura e funcionamento do aplicativo desenvolvido junto aos resultados obtidos pelo mesmo.

**PALAVRAS CHAVE:** Imagem digital, Processamento digital de imagens, Estereoscopia, Anáglifo, Paralaxe, OpenVC



## **ABSTRACT**

Our subject in this review is to provide an introductory view for 3D technologic areas, as it may work as a basis to the beginning students in this area or as a reference to advanced studies. This research is divided into two parts: a mainly one about the theoretical foundations and another one, describing the application built while our review. The foundations part presents the building process of a digital image, as well it's a brief introduction to the processing of images and kinds of 3D technologies, along the facts which have importance in its history. In the complementary part we describe in a non-technical way the structure and function of the application developed with the results obtained in it.

**KEYWORDS:** Digital image, Images digital processing, Stereoscopy, Anaglyph, Parallax, OpenVC.

# SUMÁRIO

1. INTRODUÇÃO.....	10
2. FUNDAMENTAÇÃO TEÓRICA .....	12
2.1. Imagem digital.....	12
2.2. Processamento digital de imagens.....	17
2.3. Tecnologia 3D .....	17
2.4. Tipos de tecnologia 3D.....	23
3. CRIAÇÃO DE IMAGENS 3D UTILIZANDO UM PAR DE WEBCAM.....	29
3.1. Sistema proposto .....	29
3.2. Experimento .....	36
3.3. Resultados.....	38
4. CONSIDERAÇÕES FINAIS .....	42
5. REFERENCIAS .....	43

## 1. INTRODUÇÃO

As tecnologias 3D estão ganhando ênfase em diversos segmentos na medida em que os avanços tecnológicos diminuem sua complexidade, chegando a despertar interesse em pessoas com pouco ou nenhum conhecimento no assunto.

Guilherme Erwin Hartung, orientador tecnológico no Colégio Estadual Embaixador José Bonifácio, recebeu pelo MEC o prêmio “Professores do Brasil”, o seu projeto teve como objetivo apresentar aos seus alunos de Petrópolis o “Fantástico Mundo 3D”. Os alunos por meio de oficinas de biologia, física e matemática, adquirindo fundamentos teóricos sobre o funcionamento do sistema de visão humano, propagação da luz e geometria envolvida na produção de imagens 3D. Com os conhecimentos obtidos, 15 estudantes em atividades oferecidas fora do horário das aulas e sem valer nota, montaram um site com as imagens produzidas por eles mesmos.<sup>1</sup>

No mundo dos cinemas o sucesso das tecnologias 3D trouxe de volta inúmeros filmes de grande sucesso, sendo estes relançados na versão 3D. O sucesso “O Rei Leão” foi um dos filmes relançados, este ocupava a trigésima quinta posição entre os filmes mais lucrativos, com lucro de 783 milhões de dólares. Com a versão 3D passou a ocupar a nona posição entre os mais lucrativos. Em apenas 27 dias atingiu a margem de lucro de 87.416.756 milhões de dólares.

Ainda em 2011, o lançamento já 3D dos filmes “Harry Potter e as relíquias da morte”, “Transformers – O lado oculto da lua” e “Piratas do Caribe – Navegando em águas misteriosas”, além de produções grandiosas e serem continuações de filmes já existentes, cada um deles ultrapassaram a marca de 1 bilhão de dólares nas bilheterias.

Toda a comoção que envolve as tecnologias 3D agrega importância ao estudo apresentado, este pode ser usado como base de conhecimento para iniciantes ou como referencia para estudos mais avançados.

O estudo tem como objetivo mostrar o que esta por trás das tecnologias 3D, abordando fatos que marcaram sua história e fenômenos que fazem dela uma realidade. Contemplando também suas principais técnicas de aplicação. Entre elas a técnica anáglifo que será aplicada ao sistema computacional.

---

<sup>1</sup><http://info.abril.com.br/noticias/carreira/professor-usa-tecnologia-em-aula-de-fisica-e-matematica-15102012-8.shl>

Foi abordado o problema de aplicar a tecnologia 3D anáglifo com imagens de 2 câmeras. Para isso foi realizado em um primeiro momento, uma pesquisa bibliográfica em livros, artigos e estudos científicos envolvidos na natureza do problema. Posteriormente foi desenvolvido um sistema computacional protótipo que contempla:

- a. Um módulo de interface de controle no qual gerencia os eventos do sistema computacional;
- b. Um módulo de interface que permite a visualização das imagens processadas;
- c. Uma classe que implementa a técnica anáglifo capaz de gerenciar 2 objetos da classes câmera;
- d. Uma classe de interface entre o sistema computacional e a câmera que permite capturar as imagens adquiridas pela câmera e disponibilizá-la para o processamento;

Para melhor análise e validação do sistema, foi realizado um experimento em um ambiente plano com iluminação adequada. Nesse experimento, foram utilizadas duas câmeras<sup>1</sup> comercial normal de baixo custo (capaz de adquirir até 30 quadros de imagens por segundo).

---

<sup>1</sup> Câmeras de 1.3 Megapixels - Multilaser modelo WC040

## 2. FUNDAMENTAÇÃO TEÓRICA

### 2.1. Imagem digital

Iniciamos nossa discussão sobre imagem digital com a definição de uns dos mais renomados pesquisadores no assunto.

“Uma imagem digital é uma imagem  $f(x,y)$  discretizada tanto em coordenadas espaciais quanto em brilho. Uma imagem digital pode ser considerada como sendo uma matriz cujos índices de linhas e de colunas identificam um ponto na imagem, e o correspondente valor do elemento da matriz identifica o nível de cinza naquele ponto. Os elementos dessa matriz digital são chamados de elementos da imagem, elementos da figura, “pixels”, ou “pels”, estes dois últimos, abreviações de “picture elements”(elemento de figura).” (GONZALEZ; WOODS, 2000, p. 4-5, tradução Marcondes; Fontoura)

Em poucas palavras GONZALEZ e WOODS (2000) descreve a essência de uma imagem digital. Caso esta seja a sua primeira experiência no assunto, provavelmente alguns termos como “discretizada”, “coordenadas espaciais”, “elementos da imagem”, enfim. Podem parecer um pouco quanto estranhos, mas até o final deste subtítulo você compreenderá a grandeza das palavras de GONZALEZ e WOODS (2000).

Vamos levar a discussão ao ponto em que existe a necessidade de se passar uma imagem do mundo real para o mundo digital. A ideia é relativamente simples, basta pegarmos uma câmera digital, efetuar a captura da imagem e passar esta para o computador. Mas o que esta por traz de ações tão simples?. Para entender o conceito precisamos resgatar alguns fundamentos sobre imagem natural (visão humana).

Os fundamentos do sistema de visão humana é algo extremamente complexo, o objetivo aqui é trazer para discussão somente informações suficientes para melhor compreensão dos assunto que abordaremos ainda neste subtítulo.

O olho humano possui um diâmetro de 2 a 2,5 cm, a radiação luminosa refletida dos objetos do mundo real penetram no olho a partir de uma abertura frontal na íris (pupila), passando pelo cristalino até chegar na retina, esta constitui a camada posterior do globo ocular, por onde a imagem da cena focalizada é formada. Os fotossensores, cones e bastonetes que compõem a retina, transformam a energia luminosa em sinais elétricos que são transmitidos ao cérebro, sendo processados posteriormente pelo mesmo. Enfatizando que cones e bastonetes possuem características distintas, enquanto o primeiro é excitado somente quando a cena focalizada esta suficientemente iluminada. Imagens resultantes destes são de alta resolução e coloridas. O segundo são excitados quando o nível de luz é baixo, resultando em imagens de baixa resolução e monocromáticas (tons de cinza). (EUSTÁQUIO, 2001, p. 5)

Como discutido anteriormente a imagem é formada pela luz refletida dos objetos que compõe a cena, mas como distinguimos as inúmeras cores presentes nos objetos?.

A natureza da luz apresenta um comportamento ondulado sendo caracterizado por sua frequência e comprimento de onda. As frequências eletromagnéticas a qual o sistema de visão humana é sensível estão contidas na escala de 400 a 700nm (Nanômetro), conhecida como luz visível, logo o comprimento de onda fora desta faixa não são percebidas pelo olho humano. Dentre o intervalo da faixa visível ao olho humano, distinguem-se comprimentos de onda diferentes como cores diferentes, tal que fontes de radiação com um único comprimento de onda são compreendidas como fontes de espectro puro, ou seja, a intensidade de radiação eletromagnética com relação a seu comprimento de onda, conhecida como monocromática. (EUSTÁQUIO, 2001, p. 4)

**Tabela 1 - Comprimento de onda (Faixa visível)**

Primaria	Comprimento de onda	Combinação	Cor secundaria
Vermelho	700,0 nm	Vermelho + verde	Amarelo
Verde	546,1 nm	Vermelho + azul	Magenta
Azul	435,8 nm	Verde + azul	Ciano

Fonte: (Fig. 5 – Modelo cromático RGB (EUSTÁQUIO, 2001, p. 6))

Segundo GONZALEZ e WOODS, 2000, uma imagem digital é uma imagem  $f(x,y)$  discretizada tanto em coordenadas espaciais quanto em brilho, mas uma imagem real é representada por meio de uma função contínua  $f(x)$ , logo esta é impossível de ser representada no computador, por tanto a mesma só pode ser simulada. (ESCAÑO, 1999, p. 10)

[...]“Podemos pensar em uma função contínua como uma corda sem emendas que se estende indefinidamente(ou seja infinita) e por mais que você corte a corda, sempre poderá corta-la um pouco menor indefinidamente(ou seja infinitesimal). Matematicamente estamos falando de uma função real definida na reta toda (em  $\mathbb{R}$ ), onde para cada valor de  $x$  fornecido sempre existe um único valor de  $f(x)$  obtido.”(ESCAÑO, 1999, p. 9)

Sendo assim o termo “discretizada” usado por GONZALEZ e WOODS (2000), foi no sentido de que é preciso transformar a função contínua em uma função discreta.

“O processo para trazer uma função contínua para o computador é discretizando-a (ou digitalizando-a), ou melhor, tornando valores pontuais ao longo de  $x$  e guardando o valor de  $f(x)$  correspondente (é claro que o eixo  $f(x)$  também é contínuo, assim também precisaremos discretiza-lo). O processo de discretização do eixo  $x$  (o domínio) é chamado de Amostragem, o do eixo  $f(x)$  (o contradomínio) é chamado de Quantização.”(ESCAÑO, 1999, p. 10)

Agora temos uma melhor compreensão sobre a definição de imagem digital apresentada por GONZALEZ e WOODS, mas e os novos termos inseridos a na discussão por ESCAÑO?. O que significa “Amostragem” e “Quantização” ?. A julgar pelo nome parece algo complexo, mas só parece.

Amostragem é o grau de detalhes preceptivos da imagem, ou seja, a sua resolução espacial ou dimensão da matriz  $f(x,y)$  que representa a imagem. Quantização é a resolução de níveis de cinza (ou brilho), ou seja, o intervalo de níveis de cinza que um ponto da imagem  $f(x,y)$  pode assumir. (CASTRO, 2006, p. 2-3)

Antes de finalizar a discussão sobre imagem digital, vamos complementar o conceito de função contínua e função discreta com o paradigma dos quatro universos.

“O Paradigma especifica 4 universos. O universo Físico, onde estão os objetos do mundo real. O universo Matemático, onde formulamos descrições abstratas desses objetos. O universo de Representação vai permitir trazer essas descrições abstratas para o mundo digital, e é onde se dará a discretização dos sinais contínuos. O universo de implementação é onde é feita a codificação do sinal discretizado na memória do computador através de uma estrutura de dados. (ESCAÑO, 1999, p. 10)

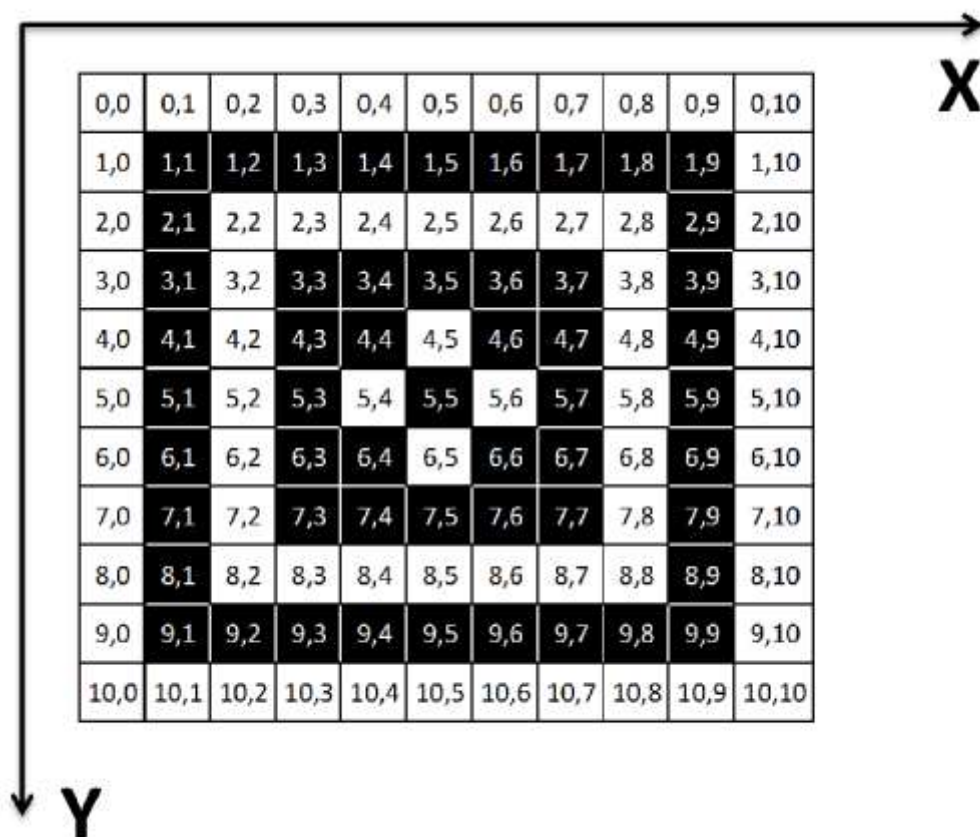


Figura 1 - Coordenadas  $f(x,y)$

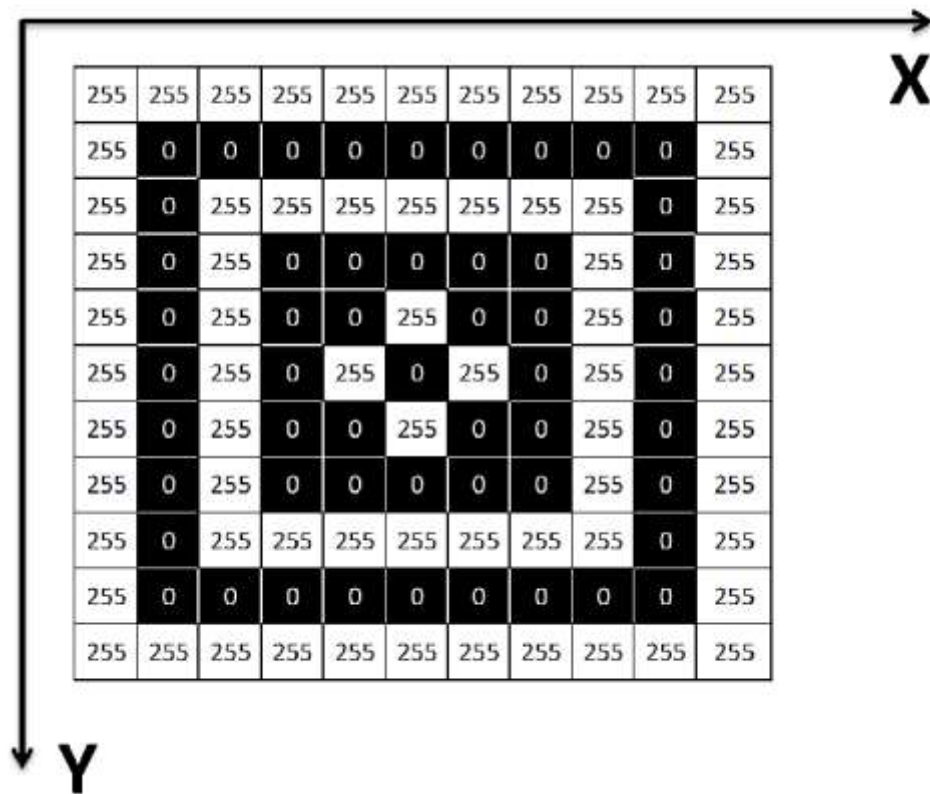


Figura 2 - Matriz  $f(x,y)$

As imagens “Figura 1 - Coordenadas  $f(x,y)$ ” e “Figura 2 - Matriz  $f(x,y)$ ” retratam todo o conceito levantado acima sobre imagem digital, onde a figura 1 ilustra a discretização em coordenadas  $f(x,y)$  e a figura 2 a quantização em uma matriz  $f(x,y)$ .



Figura 3 - Lena (monocromática)

Fonte: ([http://www.ricbit.com/uploaded\\_images/lena-713372.jpg](http://www.ricbit.com/uploaded_images/lena-713372.jpg) (transformada em monocromática))



A imagem “Figura 3 - Lena (monocromática)” exibida anteriormente tem a resolução de 220 x 220, logo sua representação digital resulta em uma matriz  $[0-255][0-255]$  com 48400 pixels. Note que a imagem não possui cores, mas sim tons de cinza (fontes de espectro puro), que podem variar de 0 a 255, sendo 0 totalmente preto e 255 totalmente branco.



**Figura 4 - Lena (RGB)**

Fonte: ([http://www.ricbit.com/uploaded\\_images/lena-713372.jpg](http://www.ricbit.com/uploaded_images/lena-713372.jpg))

As imagens “Figura 3 - Lena (monocromática)” e “Figura 4 - Lena (RGB)” referenciam a mesma cena, possuindo a mesma resolução e quantidade de pixels, mas a segunda esta é representada no modelo RGB (Red, Green, Blue), onde a intensidade das ondas eletromagnética do vermelho (700nm), verde (546,1nm) e azul (435,8nm) são representadas por matrizes distintas  $f(x,y)$  ( $R(x,y)$ ,  $G(x,y)$ ,  $B(x,y)$ ). (ANTÔNIO; CARLOS, 2011, p. 63)

Podemos agora finalizar a discussão alegando que, a representação de uma imagem real no mundo digital é feita a partir da associação desta a um modelo matemático, onde em um processo de discretização consegue-se chegar a uma estrutura de dados capaz de representar o sinal em um computador.

## 2.2. Processamento digital de imagens

Agora que já sabemos como uma imagem é representada no computador, vamos a uma breve introdução sobre processamento digital de imagens.

Processar uma imagem digital consiste em manipular os seus *pixels* e/ou o valor quantizado associado com cada *pixel*, para transformá-la numa outra imagem. Esse processo permite: remover ruídos; alterar valores relacionados com contraste, brilho, cor e saturação associada com cada pixel da imagem; aplicar transformações geométricas; reconstruir imagens corrompidas; executar processos de segmentação com a finalidade de extrair atributos específicos de elementos que compõem a imagem. (GONZALEZ; WOODS, 2000, tradução Marcondes; Fontoura)

## 2.3. Tecnologia 3D

Quando falamos sobre imagens em 3 dimensões, o popular “3D”, é comum para os não conhecedores do assunto, referente a trajetória científica ou evoluções desta tecnologia, dizerem que se trata de uma tecnologia nova, recente, onde acredita-se que um óculos 3D é o único requisito necessário para ver as imagens em 3D.

MORIN (2006) diz que: [...]“é complexo o que não pode se resumir numa palavra-chave, o que não pode ser reduzido a uma lei nem a uma ideia simples”.(MORIN, 2006, p. 5, tradução LISBOA)

Esta afirmação não estaria tão errada, pois a ideia central é relativamente simples: “ver profundidade nas imagens criadas a partir de tecnologias”. Com a popularidade atual e exploração da tecnologia 3D em diversos segmentos, logo se deduz como se chegou à ideia: “Imagem 3D é uma tecnologia nova que permite ver profundidade nas imagens”. Estudos comprovam que a ideia de criar imagens que favoreça a concepção de profundidade em observadores é muito antiga.

Os princípios e as primeiras experiências com intuito de favorecer a percepção de profundidade em imagens surgiram a mais de meio século, o primeiro filme em 3D exibido em um cinema nos Estados Unidos no ano de 1952, na época a sensação de ver as imagens saindo da tela causou muito entusiasmo sobre a tecnologia 3D. (LANDIM, 2009)

Novos avanços não muitos significativos continuaram surgindo, mas ainda havia um grande desconforto para os espectadores que os assistiam, os recursos tecnológicos necessários para exibição e observação de imagens 3D da época não se compara ao que se tem hoje, era preciso melhorar o som, o formato de exibição de imagens, as salas de cinema e os óculos utilizados pelos espectadores, que mesmo sendo de papel eram desconfortáveis. Em algumas pessoas a experiência da percepção de profundidade em imagens tecnológicas, provocavam efeitos colaterais, como dores de cabeça e enjoo. (LANDIM, 2009)

Para WAN, ANJOS e MONTEIRO (2010) os pioneiros da tecnologia 3D foram os artistas Giotto e Carvaggio, que no século 13 já usavam essa técnica em suas pinturas.



**Figura 5 - Pintura de Giotto do século 13**

Fonte: (<http://conhecaolivreiro.com.br/pdf/livros/cultura/833809.pdf> p. 2)

Em 1611 o astrônomo Johannes Kepler publicou um artigo descrevendo a estereoscopia, onde com detalhes, a teoria da visão estérea baseada nas projeções dos raios luminosos nos olhos é descrita por ele. Tempo depois o físico Sir Charles Wheatstone, inventor do estereoscópio, aplica a técnica estereoscopia em sua invenção. Em 1848, dez anos depois do invento de Sir Charles Wheatstone, David Brewster melhorou o estereoscópio e comercializou o produto, vendendo mais de 250 mil exemplares. (WAN; ANJOS; MONTEIRO, 2010, p. 1)

O primeiro filme em 3D documentado foi exibido em 1922 em Los Angeles, realizado e produzido por Harry K. Fairall que deu o nome de “The Power of Love”, esta foi à primeira experiência de cinema 3D. (WAN; ANJOS; MONTEIRO, 2010, p. 2)

AZEVEDO (2012) concorda com WAN, ANJOS e MONTEIRO (2010) sobre o fato de que o primeiro filme em três dimensões foi exibido em 1922. Complementando ainda que na década de 50 foi criada a Kinematoscope, câmera de animação de som. Nesse período as pessoas estavam deixando de frequentar as salas de cinema devido à popularização da televisão. Os produtores dos estúdios decidiram assim criar algo para fazer com que o público voltassem às salas de projeção. Foi quando surgiu a tecnologia Naturalvision, formada em uma imagem anáglifa que passando a ter efeitos estereoscópico (tridimensional), quando visualizada por um óculos composto de lentes plásticas nas cores azul e vermelha. (AZEVEDO, 2012, p. 2-3)

O filme “BwanaDevil” foi o primeiro a apresentar a tecnologia anáglifo, seu anúncio de divulgação: “O que você quer? Um bom filme ou um leão no seu colo?” ficou marcado na história. (AZEVEDO, 2012, p. 3)

AZEVEDO (2012) e WAN, ANJOS e MONTEIRO (2010) discordam de LANDIM (2009) a respeito do primeiro filme em três dimensões a ser exibido, e que com ele surgia a os fundamentos para tecnologia 3D, mas todos estão de acordo do sucesso provocado pela exibição do filme “BwanaDevil”, das limitações existentes na época e que a o sucesso da tecnologia 3D existente até o momento ainda dependem da estereoscopia.

Mas o que seria “estereoscopia”? Antes de trazer os seus fundamentos a discussão, vamos retratar alguns dos aspectos relacionados ao processamento de imagens ocorrentes no sistema de visão humano, complementando discussões de subtítulos anteriores e direcionando os conhecimentos obtidos até aqui para a compreensão da discussão atual “Tecnologia 3D”.

Muitas vezes somos surpreendidos por inovações sofisticadas nunca antes imaginadas que saíam do estágio de sonho, principalmente tecnologias voltadas para o processamento digital de imagens.

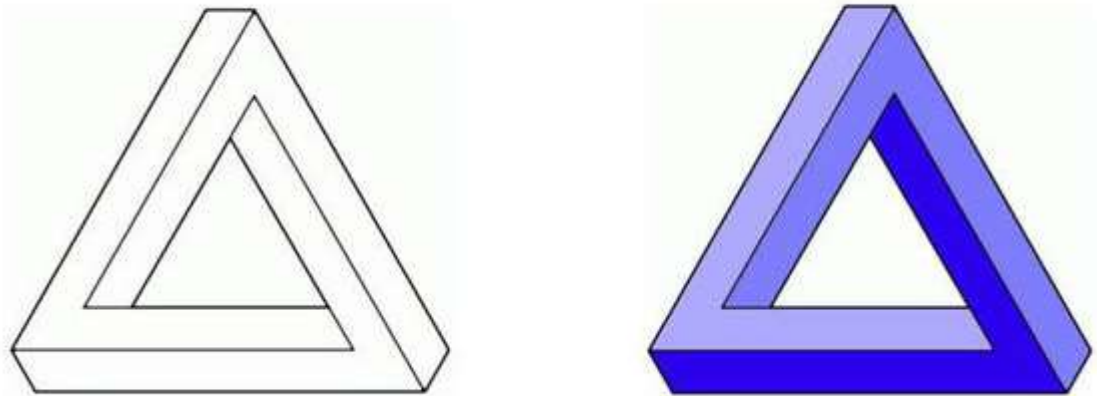
Quando falamos em novas tecnologias voltadas para o processamento digital de imagens, em caráter de software e hardware, nada existente até o momento se compara ao sistema de visão humana em termos de velocidade, qualidade e flexibilidade em adaptação.

O simples fato de você abrir os olhos acarreta um volume de processamento muito intenso, mas como a capacidade de processamento do cérebro é muito alta o processo se torna imperceptível. (HOFFMAN, 2000, p. 1, tradução CABRAL)

HOFFMAN (2000) ilustra o conceito do gênio criativo e sua inteligência visual, partindo do princípio que tudo que vemos: cor, sombra, textura, movimento, formas, objetos visuais e cenas visuais completas. É fruto de nossa criação.

”Você é um gênio criativo. Seu gênio criativo é tão completo que ele parece, para você e para os outros, algo que se realiza sem esforço. No entanto, ele supera os esforços mais valorosos dos mais velozes supercomputadores atuais. Para invocá-lo, você só necessita abrir seus olhos. Isso pode soar como o mantra de uma nova terapia, ou os dizeres do papelzinho que acompanha o biscoito da sorte chinês. Mas, ao contrário, esta é a conclusão fundamentada de pesquisadores no campo da ciência cognitiva. O que ocorre quando você vê não é um processo de estímulos e respostas sem participação da mente, como pensavam os behavioristas durante boa parte do século XX, mas um processo sofisticado de construção cujas complicações estamos começando a compreender atualmente. Numa fração de segundo, sua inteligência visual pode construir a pavonada e as cores de um pavão, ou a corrida graciosa de um leopardo, as cores flamejantes de um crepúsculo no oceano, ou as nuances de luz numa floresta ao anoitecer, ou qualquer das inúmeras outras cenas de tal sutileza e complexidade. Você ultrapassa de longe o maior dos pintores, em repertório e velocidade.”(HOFFMAN, 2000, p. 1, tradução CABRAL)

A forma como os objetos são criados pelo sistema de visão é realmente fantástico e curioso ao mesmo tempo. Independente dos ambientes ou objetos observados, conhecidos ou não conhecidos. O sistema de visão constrói estes de forma lógica baseado em seu conjunto de regras, onde não existem exceções.



**Figura 6 - Triângulo impossível**

Fonte: (<http://i8.photobucket.com/albums/a40/designiade/construodotrinaguloimpossivel.jpg>)

A imagem cima “Figura 6 - Triângulo impossível” foi inventada por Oscar Reutervard em 1934. Ao observar esta imagem percebemos algo curioso, mesmo sabendo que esta é uma figura em 2 dimensões, o sistema de visão interpreta como sendo de 3 dimensões. Essa interpretação não é por acaso, mas sim o resultado de um conjunto de regras lógicas para interpretação de imagens. (HOFFMAN, 2000, p. 27-165, tradução CABRAL)

Há princípio pode parecer estranho ou até mesmo uma falha no sistema de visão, pois temos a certeza que a figura não possui profundidade, mas mesmo assim não conseguimos interpretar esta como tal. A esse pronto compreendemos que não temos o controle sobre as imagens construídas pelo sistema de visão, e que nem tudo que vemos é representação do real. Parece algo perigoso e falho, mas só parece. Nosso julgamento final sobre o que é consistente ou inconsistente não depende somente da imagem do ambiente ou objeto observado, mas sim de toda a experiência visual construída desde que abrimos os olhos pela primeira vez. (HOFFMAN, 2000, p. 4-5, tradução CABRAL)

“Dentre os fatos mais surpreendentes sobre a visão está o de que as crianças são gênios completos para a visão antes mesmo de aprender a andar. Antes de um ano de idade, podem construir um mundo visual em três dimensões, navegar através dele de maneira bastante orientada, mesmo engatinhando, organizá-lo em objetos e agarrar, morder e reconhecer esses objetos.”[...] (HOFFMAN, 2000, p. 12, tradução CABRAL)

A experiência visual é construída praticamente da mesma forma para todos os seres humanos normais, logo em casos como o Triângulo impossível, ilustrado anteriormente, existe um estímulo a criação de profundidade intencional sobre a imagem observada, ou seja, tal imagem não existe na natureza, elas são criadas por estudiosos da visão, com objetivo de estimular o nosso sistema de visão a criar profundidade onde de fato sabemos que não as possuem. (HOFFMAN, 2000, p. 1 - 46, tradução CABRAL)

Agora nossa discussão parou em um ponto muito interessante, pois já podemos dizer que imagem 3D não é uma tecnologia nova, e que nesta tecnologia em primeira ordem, o cérebro é o “fator x” para o seu sucesso.

Vamos agora ao temor “estereoscopia”.

“A estereoscopia é uma reprodução artificial da visão humana, que é possível pela existência de células fotossensíveis na retina. A forma fisiológica da visão humana é tridimensional. Isto ocorre porque a visão binocular, resultante da posição alinhada dos olhos, permite que um objeto seja visto de ângulos ligeiramente diferentes, um mais à direita e outro mais à esquerda. O processamento destas informações pelo sistema nervoso central propicia a percepção de profundidade. Quando já houve contato com o objeto, há uma noção imediata de tridimensionalidade, pela evocação de dados contidos na memória.”[...] (SOUSA et al., 2002, p. 769-770)

Com a definição de estereoscopia inserida na discussão por SOUSA (2002), juntamente com as discussões anteriores, compreendemos que a tecnologia 3D estimula o sistema de visão humano a criar profundidade em imagens tecnológicas, ou seja, naturalmente podemos ver o que vamos chamar aqui de “visão 3D real”, e com os estímulos da tecnologia 3D podemos ver o que chamamos aqui de “visão 3D ilusória”. A primeira quando observamos cenas naturais como: montanhas, cidades, avenidas, moveis, etc. Já na segunda quando observemos imagens criadas pelo homem, onde são aplicadas técnicas estereoscópicas.

## 2.4. Tipos de tecnologia 3D

Quando falamos em tipos de tecnologias 3D chegamos a 3 técnicas distintas: estereoscopia anaglífica, estereoscopia polarizada e estereoscopia ativa. Existem ainda algumas variações destas, mas não abordaremos neste estudo. (RAPOSO et al., 2004)

Para dar início a discussão vamos a algumas considerações sobre o termo “Paralaxe”. Este é utilizado na estereoscopia onde indica o deslocamento de um objeto ao mudar o ponto de observação, ou seja, como vimos na discussão anterior, a estereoscopia é formada por duas imagens com ângulos ligeiramente diferentes, logo paralaxe é o ponto de colisão onde a imagem da esquerda colide com a imagem da direita. (RAPOSO et al., 2004)

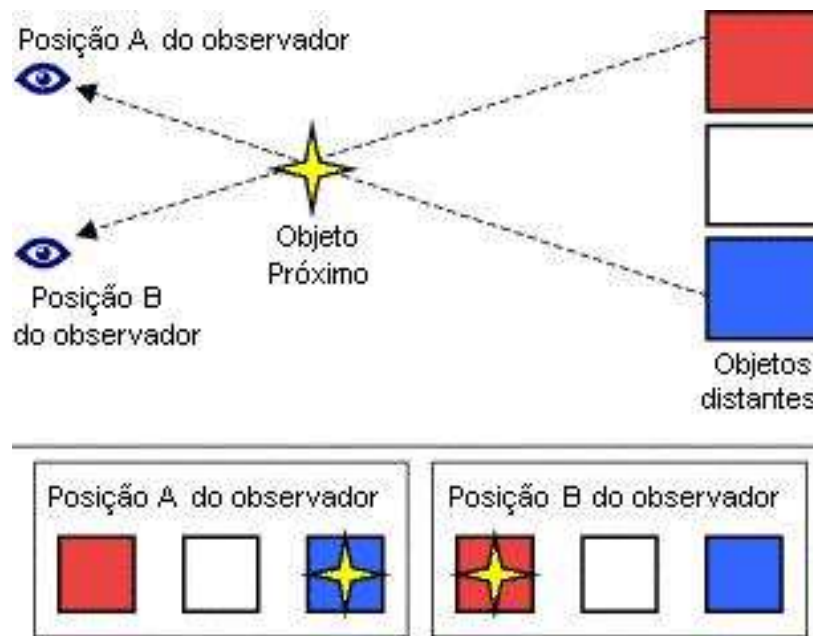


Figura 7 – Paralaxe

Fonte: (<http://astronomia.blog.br/wp-content/uploads/2006/12/200612artigofigura1.jpg>)

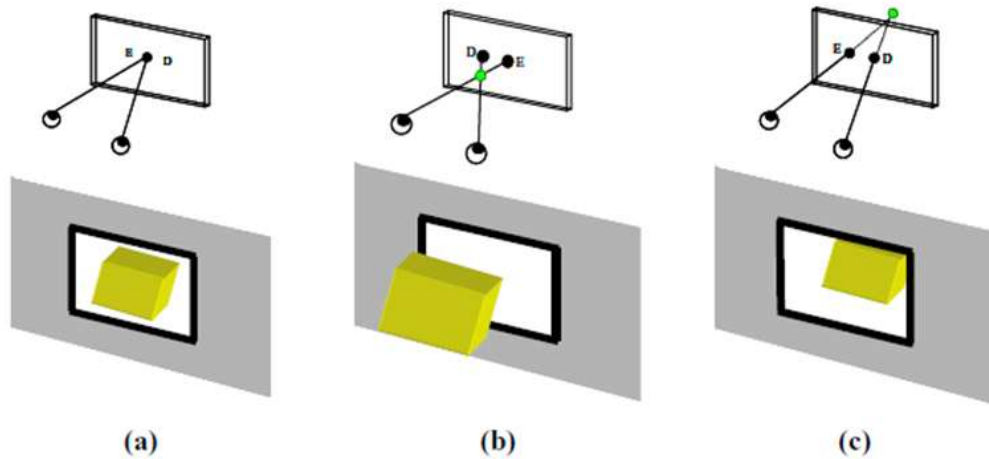
Como pode ser observado na imagem acima, tendo o controle da paralaxe, obtém-se o controle dos efeitos estereoscópicos. Quando a paralaxe é negativa<sup>1</sup>, o efeito estereoscópico provocado é a sensação de os objetos estarem saindo da tela. Paralaxe zero<sup>2</sup>, o efeito estereoscópico é menos perceptível. A paralaxe positiva<sup>3</sup> provoca um efeito contrario a paralaxe negativa, dando a sensação de os objetos estarem atrás da tela. (RAPOSO et al., 2004)

<sup>1</sup> A Paralaxe é negativa quando o ponto de colisão esta entre a tela e os olhos.

<sup>2</sup> A Paralaxe é zero quando o ponto de colisão esta junto à tela.

<sup>3</sup> A Paralaxe é positiva quando o ponto de colisão esta atrás da tela.





Tipos de paralaxe: a) Paralaxe zero (ZPS), b) Paralaxe negativa e c) Paralaxe positiva.

Figura 8 - Tipos Paralaxe

[http://www.photon3d.com.br/CSS/Imagens/paralaxe/paralaxe\\_ngtv\\_pstv.jpg](http://www.photon3d.com.br/CSS/Imagens/paralaxe/paralaxe_ngtv_pstv.jpg)

Vamos agora à técnica estereoscopia ativa. Nesse tipo de estereoscopia os óculos<sup>1</sup> são constituídos por duas lentes de cristais ou LCD. As lentes são capazes de fechar a visão dos olhos de acordo com o sinal de sincronismo emitido pelo projetor, garantindo que a imagem projetada chegue a apenas um olho por vez. O processo é muito rápido, o sistema de visão interpreta as duas imagens transmitidas individualmente para cada olho como sendo apenas uma, possuindo ângulos de visão diferentes. (RAPOSO et al., 2004)

<sup>1</sup> Os óculos 3D utilizados na estereoscopia ativa são chamados de obturadores.

## Esquema de Visualização por Obturação Estereoscopia Ativa

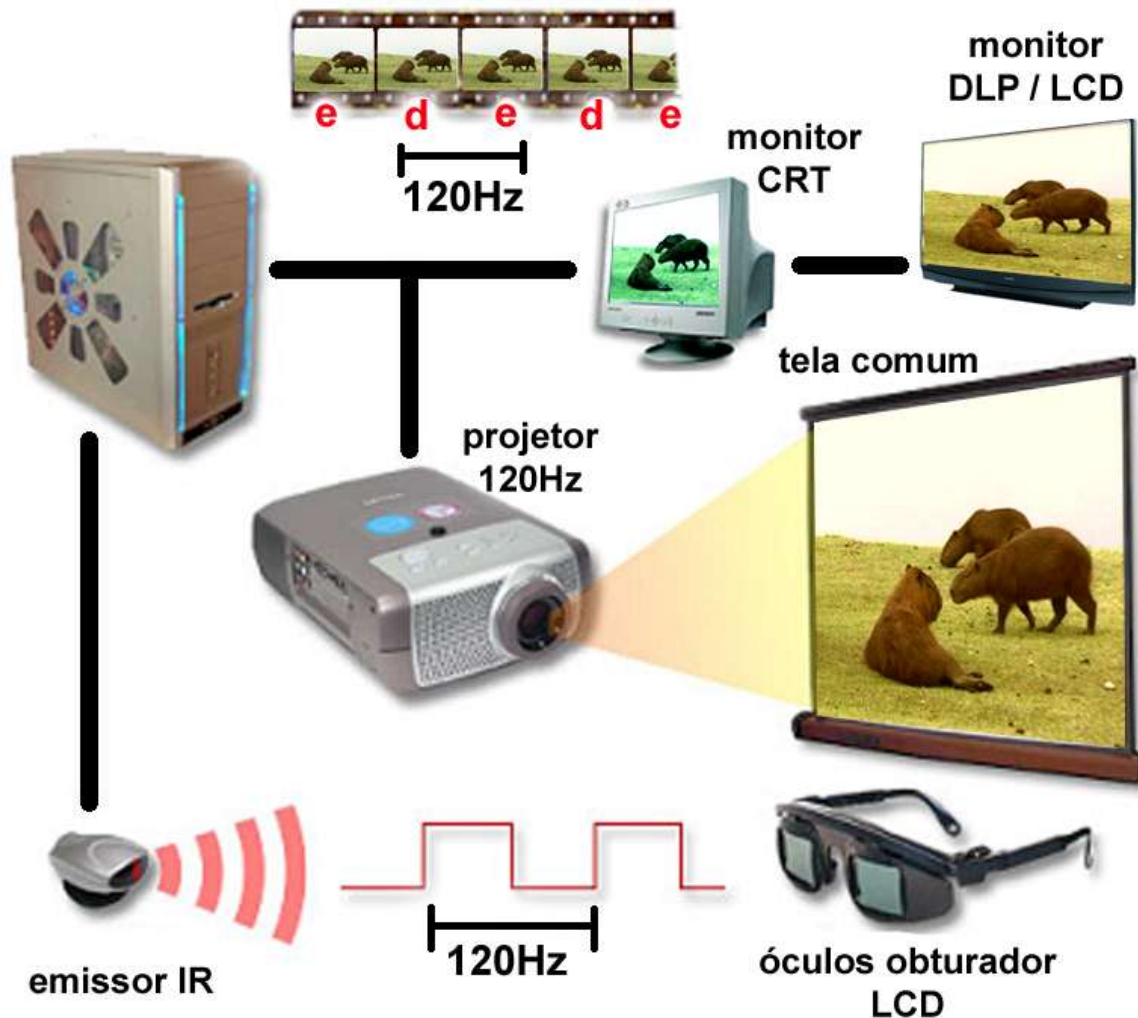


Figura 9 - Estereoscopia ativa

Fonte: ([http://www.photon3d.com.br/CSS/Imagens/3dComOculos/ESQUEMA\\_POLARIZACAO\\_ATIVA\\_site.jpg](http://www.photon3d.com.br/CSS/Imagens/3dComOculos/ESQUEMA_POLARIZACAO_ATIVA_site.jpg))

Agora vamos à técnica estereoscopia polarizada. Esta técnica é a mais utilizada atualmente. Em uma tela magnetizada são transmitidas duas imagens com ângulos ligeiramente diferentes, mas cada uma das imagens recebe polarização diferente, de forma que estas são filtradas pelos óculos 3D de acordo com a sua polarização, podendo ser linear ou circular. Permite que imagens de alta qualidade sejam projetadas sem perda cores como é o caso da estereoscopia anaglífico que será apresentada posteriormente. (RAPOSO et al., 2004)

## Esquema de Projeção por Polarização Estereoscopia Passiva

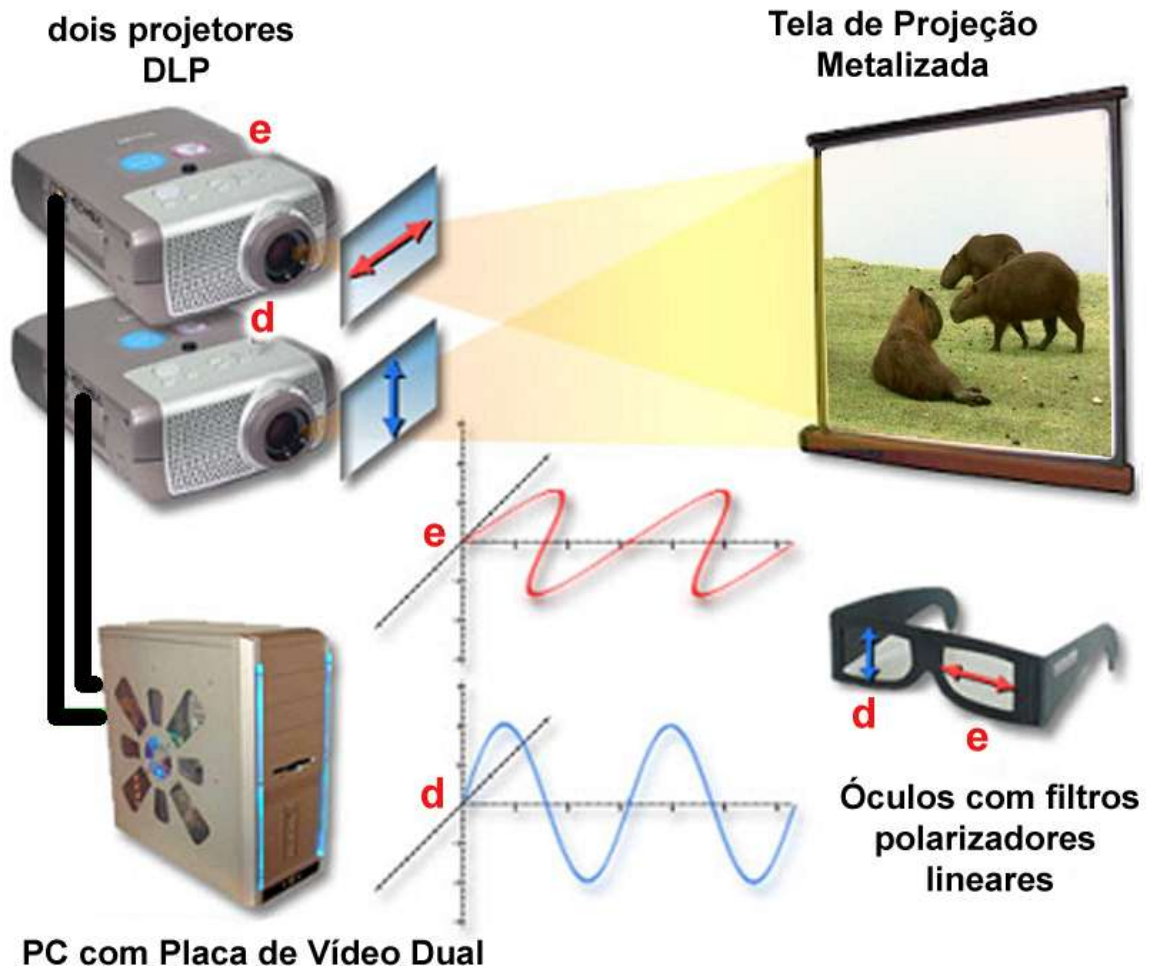


Figura 10 - Estereoscopia polarizada

Fonte: ([http://www.photon3d.com.br/CSS/Imagens/3dComOculos/ESQUEMA\\_POLARIZACAO\\_PASSIVA\\_site.jpg](http://www.photon3d.com.br/CSS/Imagens/3dComOculos/ESQUEMA_POLARIZACAO_PASSIVA_site.jpg))

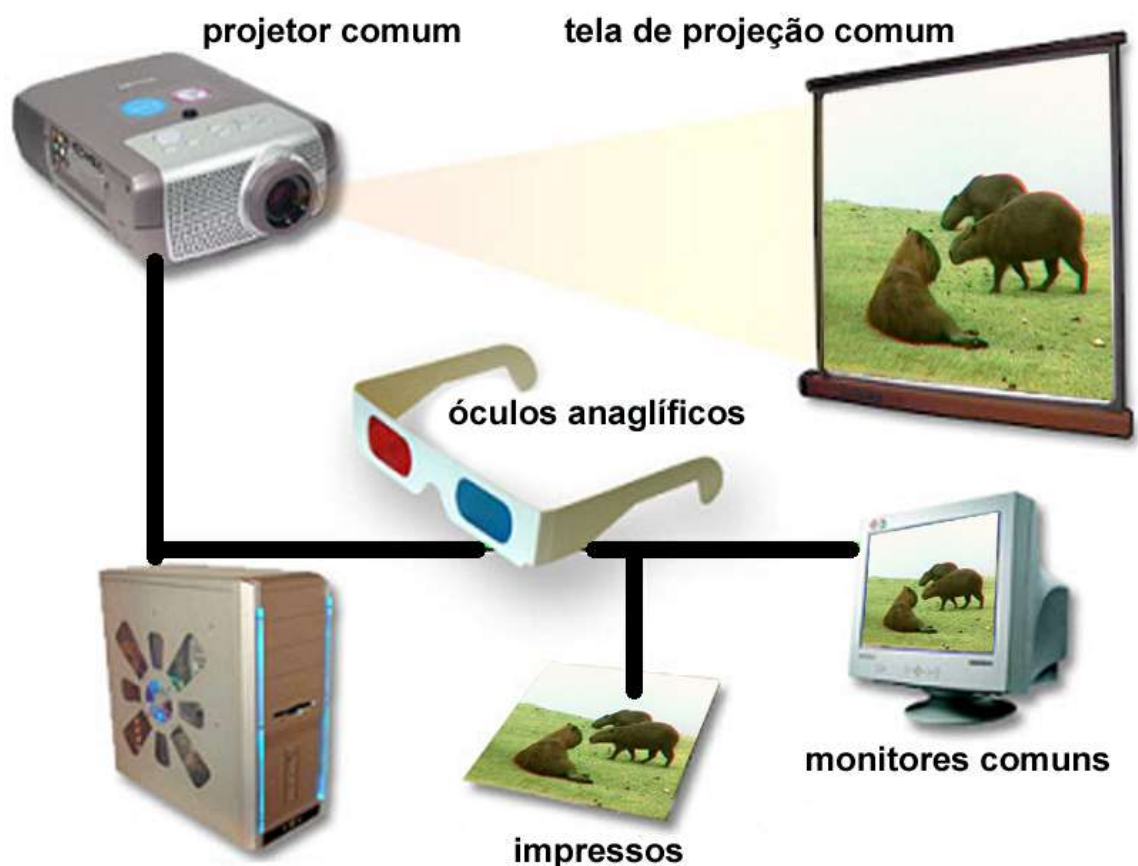
Por último abordaremos a técnica estereoscopia anáglifo. Esta sendo a nossa tecnologia de interesse. Como foi levantado anteriormente, ao aplicar esta técnica em uma imagem de alta qualidade, a uma perda significativa de qualidade, isso devido ao seu funcionamento.

“O estéreo anáglifo (do inglês anaglyph) é o exemplo mais comum e muito usado nos cinemas antigamente. Neste estéreo, a filtragem é feita por cores. No início da utilização desse estéreo, eram utilizadas duas cores apenas, vermelho e azul. Isto é, a imagem para o olho esquerdo tinha apenas componente de cor vermelha e a imagem para o olho direito componente azul.[...]” (RAPOSO et al., 2004)

Quando a técnica anáglifo é aplicada em imagens “vivas” cheia de combinações de cores, esta passa a ser representada por apenas duas, o vermelho e o azul como definido por RAPOSO et al (2004).

Atualmente o componente verde é adicionado a imagem anáglifo, com isso se consegue uma melhora na qualidade da imagem, passando a ter 3 canais de cores. Os óculos utilizados nesta técnica podem ser confeccionados manualmente, pois as lentes utilizadas nos óculos 3D podem ser feitas com papel celofane, onde o vermelho para o olho esquerdo e azul para o olho direito. Logo ao se observar uma imagem anáglifo a cor vermelha é filtrada pela lente vermelha, e a cor azul é filtrada pela lente azul, fazendo com que cada olho receba imagens diferentes, sendo estas constituídas pela cor filtrada mais a cor verde. (RAPOSO et al., 2004)

### **Esquema de Visualização Anaglífica (separação cromática)**



**Figura 11 - Estereoscopia anáglifo**

Fonte:

([http://www.photon3d.com.br/CSS/Imagens/3dComOculos/ESQUEMA\\_ESTEREOSCOPIA\\_ANAGL%C3%8DFICA\\_SITE.jpg](http://www.photon3d.com.br/CSS/Imagens/3dComOculos/ESQUEMA_ESTEREOSCOPIA_ANAGL%C3%8DFICA_SITE.jpg))

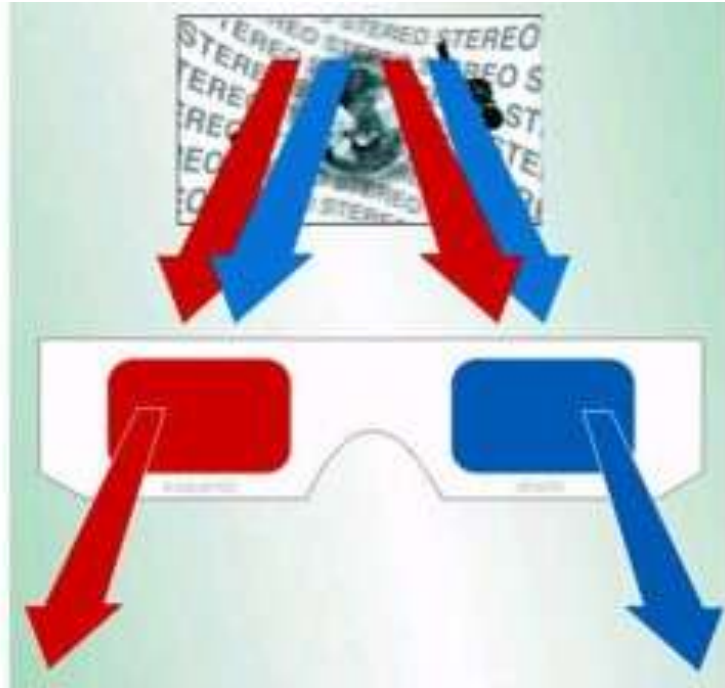


Figura 12 - Óculos Anáglifo

Fonte: (<http://www.blogintellectus.com.br/fisica/wp-content/uploads/2010/04/3d-cores-300x225.jpg>)

### 3. CRIAÇÃO DE IMAGENS 3D UTILIZANDO UM PAR DE WEBCAM

#### 3.1. Sistema proposto

O sistema desenvolvido durante o estudo apresentado, tem como objetivo aplicar a técnica anáglifo utilizando imagens capturadas por duas webcam's em tempo de real, portando uma interface amigável de auto compreensão.

Como linguagem de programação foi escolhida o C#, variáveis como tempo disponível para desenvolvimento, nível de complexidade que se pretendia alcançar e nível de recursos presentes na IDE de desenvolvimento Visual Studio 2010, foram os fatores cruciais na escolha pela linguagem C#.

O desenvolvimento foi dividido em 3 etapas, pesquisa, desenvolvimento de protótipos, e desenvolvimento final.

Na primeira etapa, foi encontrado um grande desafio, capturar as imagens das duas webcam. Pesquisas foram feitas de modo a explorar os recursos existentes no Microsoft Net. Framework 4.0<sup>1</sup> para o processamento de imagens. A pesquisa foi interrompida, pois foi encontrado uma biblioteca para processamento de imagens, conhecida como OpenCV.

“A biblioteca OpenCV foi desenvolvida pela Intel e possui mais de 500 funções. Foi idealizada com o objetivo de tornar a visão computacional acessível a usuários e programadores em áreas tais como a interação humana-computador em tempo real e a robótica. A biblioteca esta disponível com o código fonte e os executáveis (binários) otimizados para os processadores Intel. Um programa OpenCV, ao ser executados, invoca automaticamente uma DLL (DynamicLinked Library) que detecta o tipo de processador e carrega, por sua vez, a DLL otimizada para este processador. Juntamente com o pacote OpenCV é oferecida a biblioteca IPL (ImageProcessing Library), Da qual a OpenCV depende parcialmente, além de documentação e um conjunto de códigos exemplos” (ANTÔNIO; CARLOS, 2011, pag115)

Originalmente a biblioteca OpenCV é escrita em linguagem C++, mas existem versões traduzidas em diversas linguagens, no sistema computacional foi utilizado a distribuição Emgucv C# versão 2.3.0.1416, para plataforma Windows 64 bits.

Agora iniciou a segunda etapa do desenvolvimento. Estudar a documentação e os códigos fontes de exemplos para o desenvolvimento de protótipos. No estudo não foi encontrado uma classe que aplicasse a técnica anáglifo a partir de duas imagens, com o comportamento igual ao esperado. Mas foi encontrado na biblioteca 4 classes que juntamente com uma estrutura de dados, conseguiria alcançar o objetivo do sistema.

---

<sup>1</sup> Base do desenvolvimento utilizando a IDE Visual Studio 2010

Tabela 2 - Classes Emguv utilizadas

Classe	Função
Capture() ou Capture(0), Capture(1)	Instancia o objeto da classe Capture: este possui métodos e propriedades equivalente a câmera, 0 e 1 são equivalentes a câmeras instaladas no computador, onde 0 = câmera 1 e 1 = câmera 2(em caso de uma terceira câmera 2 = câmera 3)
Image<Bgr, byte>	Classe que representa uma imagem digital neste caso aplicando o modelo RGB.
Image<Gray, byte>	Classe que representa uma imagem digital neste caso aplicando o modelo Monocromático.
ImageViewer	Classe que contem métodos e propriedades com intuito de exibir a imagem digital na tela.

Com base nessas 4 classes exibidas acima, e a experiência obtida pela aplicação das mesmas em protótipos, foi idealizada a seguinte estrutura de dados: Interface gráfica para os controles, interface gráfica para exibição de imagem, uma classe câmera onde contempla as 4 classes da biblioteca Emguv e uma classe câmera 3D que contempla 2 objetos da classe câmera e métodos para manipulação destas.

Tabela 3 - Estrutura de dados simplificada

Nível	Classe	Função
1	Interface controles	Gerenciar as ações do objeto da classe câmera 3D
2	Interface exibição de imagens	Exibir as imagens transmitidas pelo objeto da classe câmera 3D
4	Câmera 3D	Gerenciar dois objetos da classe câmera
5	Câmera	Representar uma webcam

A classe câmera em uma visão geral é a mais complexa, pois nela contempla o maior numero de parametrizações e procedimentos. Através do seu construtor é parametrizado o id de uma das câmeras instaladas no computador, sendo o id 0 para câmera1 e id 1 para câmera2, neste momento o objeto do tipo câmera esta pronto para receber ordens como: Abrir câmera<sup>1</sup>, Fechar câmera<sup>2</sup>, Abrir visualizador<sup>3</sup> e Fechar visualizador<sup>4</sup>.

O núcleo de processamento do objeto da classe câmera quando iniciado, passa a ser processado em segundo plano<sup>5</sup>, possuindo ciclos<sup>6</sup> regulares de atividades, sempre obedecendo às parametrizações atribuídas ao objeto. Por exemplo, não se tem controle absoluto sobre o núcleo de processamento, parametrizações feitas ao objeto só terão efeito na execução do próximo ciclo, sendo assim, quando o objeto recebe uma parametrização para finalizar a captura de imagens, esta pode ser finalizada imediatamente ou apos a execução do próximo ciclo.

---

<sup>1</sup> Inicia a capturar imagens.

<sup>2</sup> Finaliza a captura de imagens.

<sup>3</sup> Abre uma tela gráfica onde é exibida a imagem processada pelo núcleo de processamento do objeto tipo câmera.

<sup>4</sup> Fecha a tela gráfica onde é exibida a imagem processada pelo núcleo.

<sup>5</sup> Em thread diferente; Execução em paralelo.

<sup>6</sup> O tempo de vida de um ciclo é de menos de 0.5 segundos



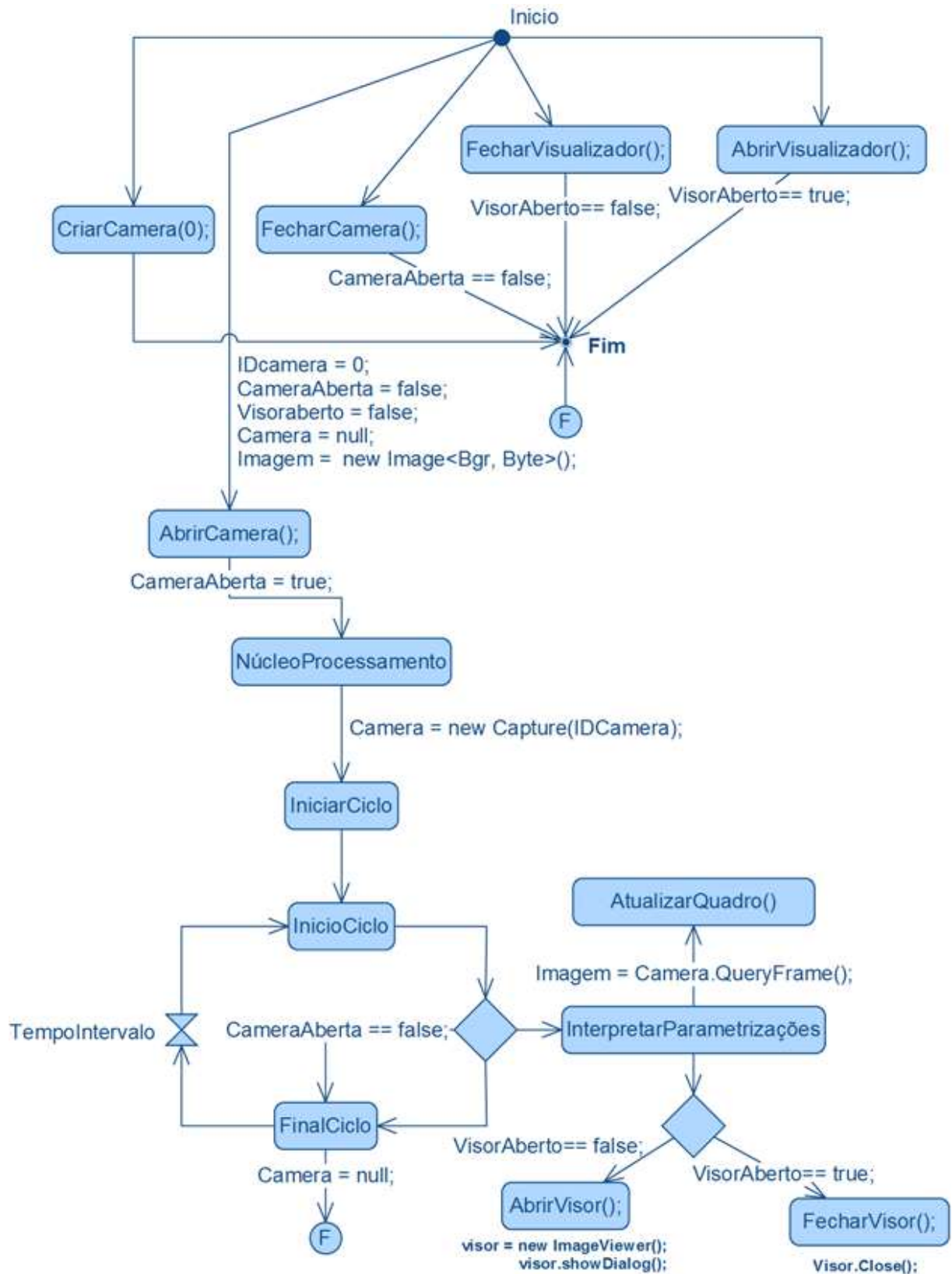


Figura 13 - Visão simplificada da classe Câmera (câmera1 id = 0)

A classe câmera 3D é responsável pelo bom funcionamento do sistema, pois gerência dois objetos da classe câmera, que são processados em segundo plano, juntamente com o seu próprio núcleo de processamento que assim como o da classe câmera, também é processado de forma paralela de acordo com as parametrizações inseridas no objeto. Temos então o seguinte cenário: 3 processos em segundo plano, onde 2 do tipo câmera e 1 do tipo câmera 3D, sendo o ultimo responsável em gerenciar os dois primeiros. Esta é a parte mais delicada do processo, mas o resultado foi motivador, não havendo travamentos ou lentidão durante o processamento das imagens.

A técnica anáglifo é aplicada pelo núcleo de processamento da classe câmera 3D, onde são fundidas as imagens dos objetos câmera 1 e 2, o primeiro quando em modo 3D<sup>1</sup> contém a imagem da banda vermelha (câmera1)<sup>2</sup> e o segundo contem a imagem das bandas verdes e azul (câmera2)<sup>3</sup>.

---

<sup>1</sup>Exibir3D = true.

<sup>2</sup>camera1[2] matriz f(x,y) canal vermelho.

<sup>3</sup>Camera2[0] matriz f(x,y) canal azul; Camera2[1] matriz f(x,y) canal verde.

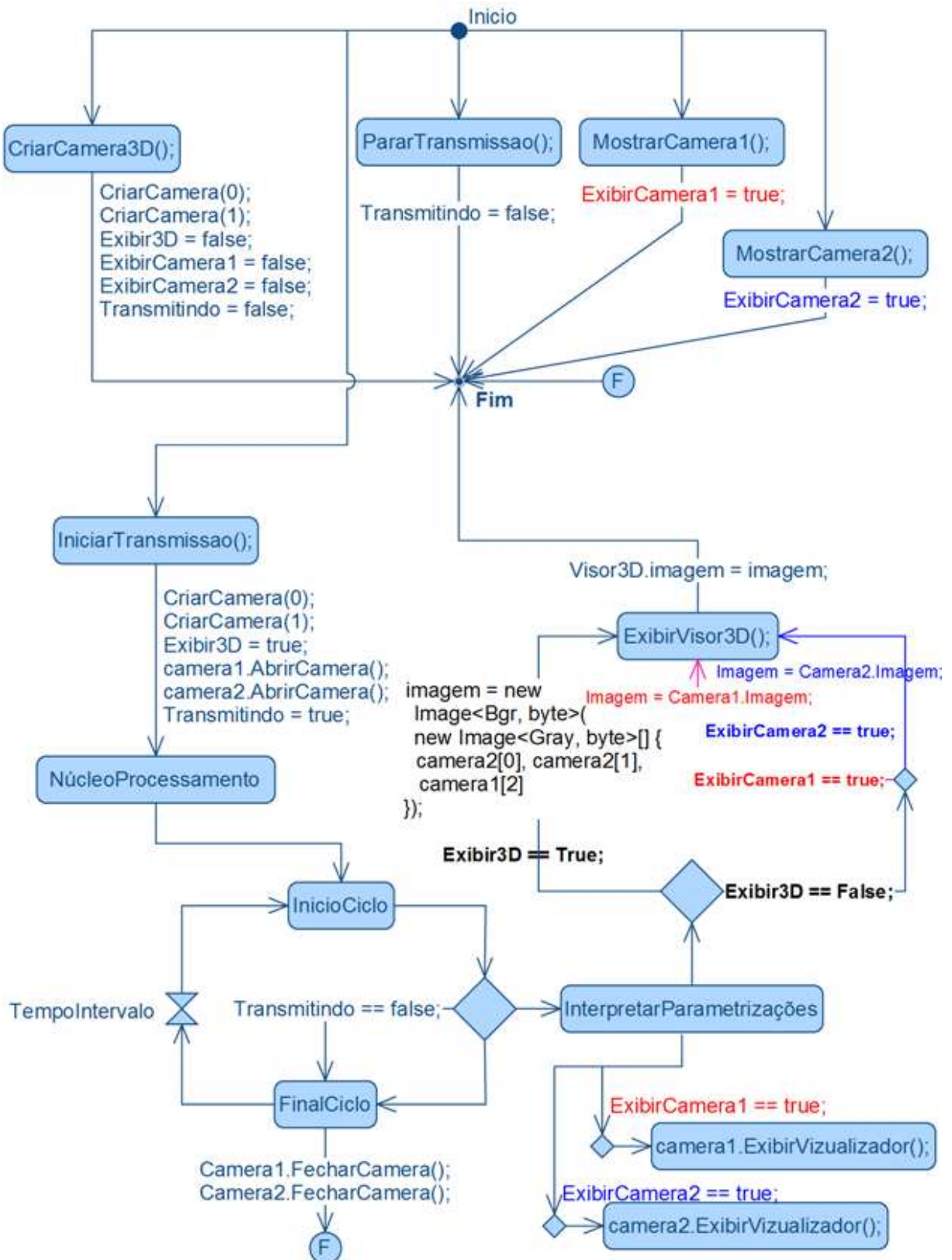
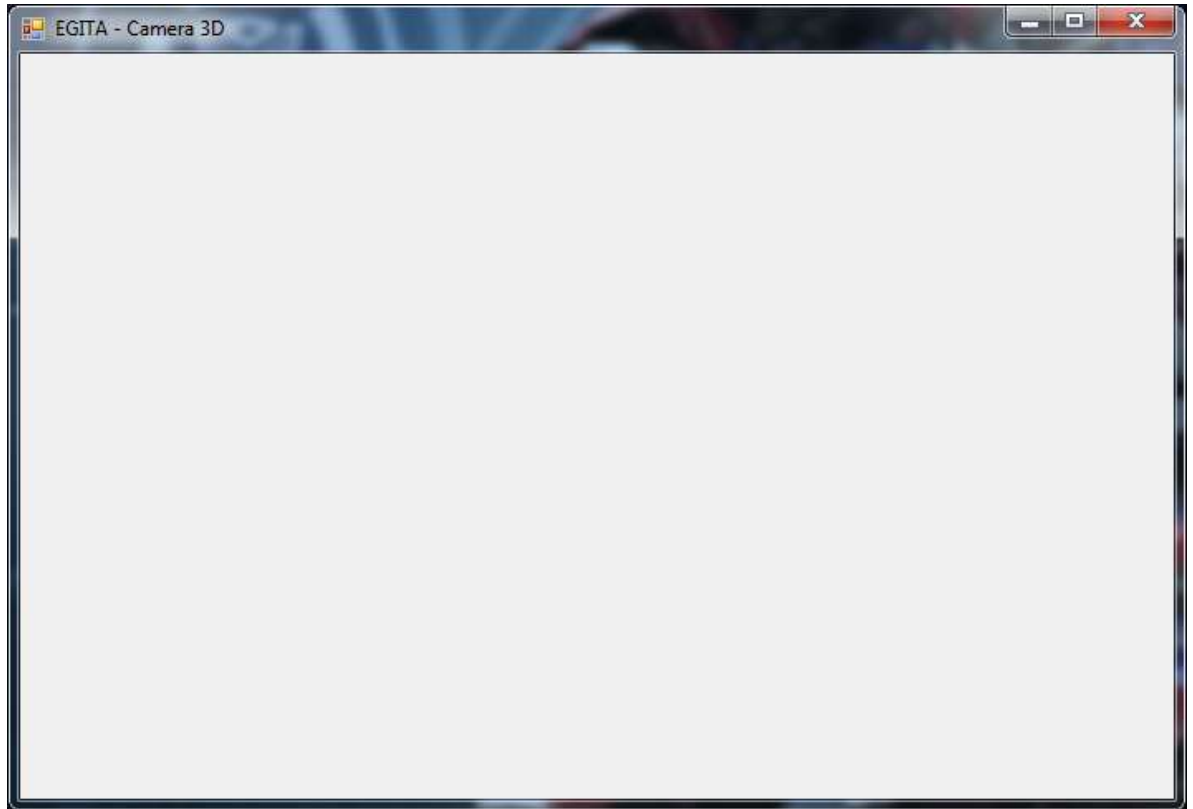


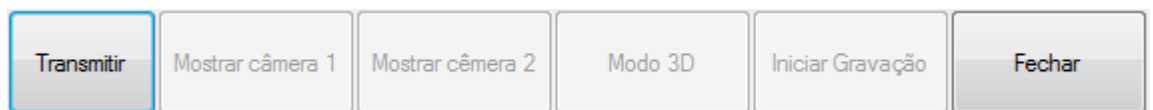
Figura 14 - Visão simplificada da classe Câmera 3D

A Interface “exibição de imagens” exibe as imagens transmitidas pelo objeto da classe 3D.

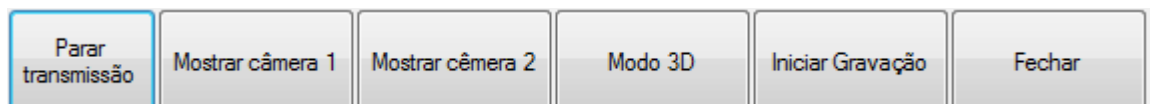


**Figura 15 - Visor da Câmera 3D**

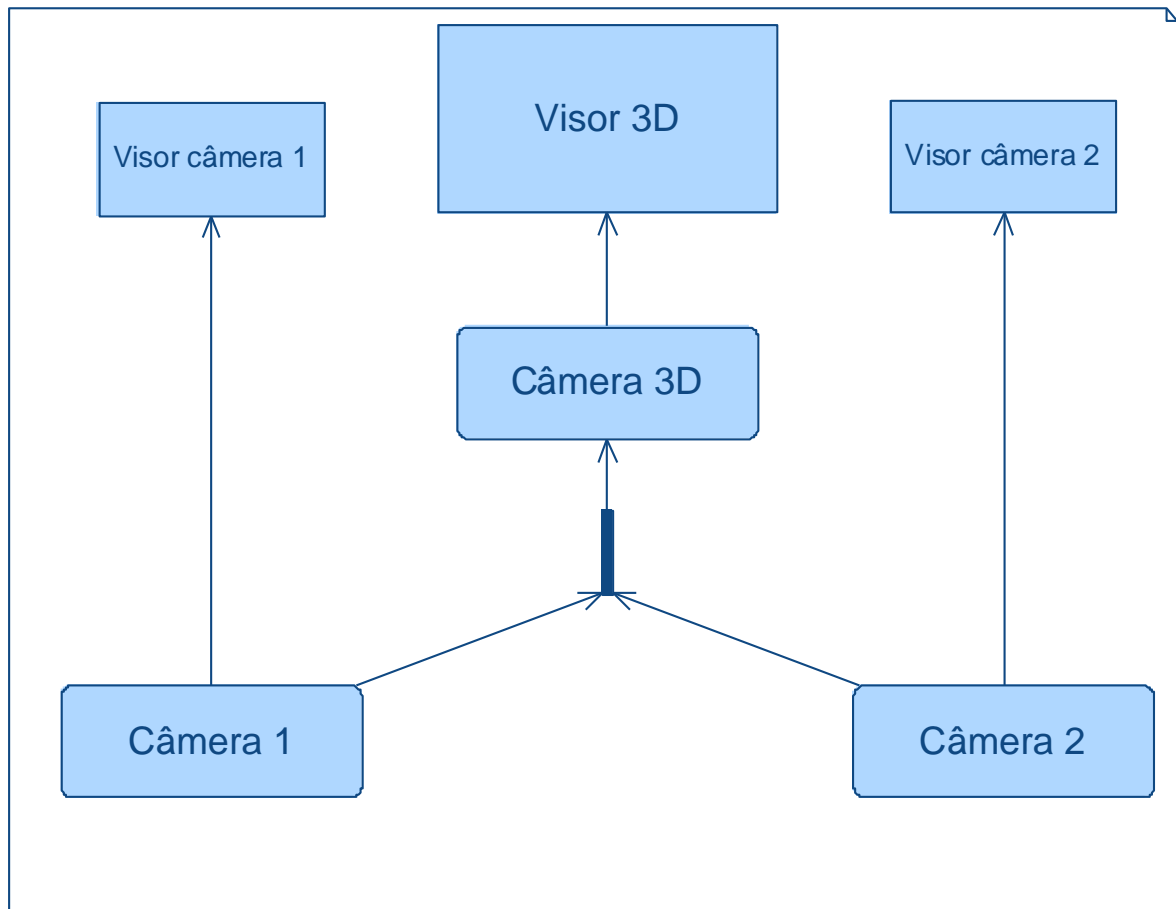
A Interface de controles apresenta de forma amigável as parametrizações disponíveis para o objeto da classe câmera 3D e fechamento do sistema.



**Figura 16 - Controles em estado inicial**



**Figura 17 - Controles em estado de transmissão**



**Figura 18 - Visão geral simplificada**

### 3.2. Experimento

O experimento iniciou-se com a preparação da cena, onde uma área de 98cm x 1,10cm foi revestida com uma camada de papel sulfite, evitando assim interferências e reflexos a partir do piso brilhante.

As câmeras foram presas juntas a um separador, de forma que a distância entre o centro de suas lentes ficou em 3cm. As câmeras foram posicionadas a 72,5cm acima da cena montada. Embora as duas câmeras fossem do mesmo modelo e fabricante, apresentaram diferenças de foco e ângulo de visão interna, depois de algumas tentativas chegou-se a um alinhamento aceitável, mantendo ainda a mesma distância entre o centro das lentes.

Feito a montagem da cena e com as câmeras ajustadas, iniciou-se a coleta de imagens geradas pelo sistema computacional.



Figura 19 - Materiais utilizados no experimento



Figura 20 - Câmeras já ajustadas



### 3.3. Resultados



Figura 21 - Livro distancia 30cm



Figura 22 - Livro distancia 103cm



Figura 23 - Livro aberto distancia 74cm

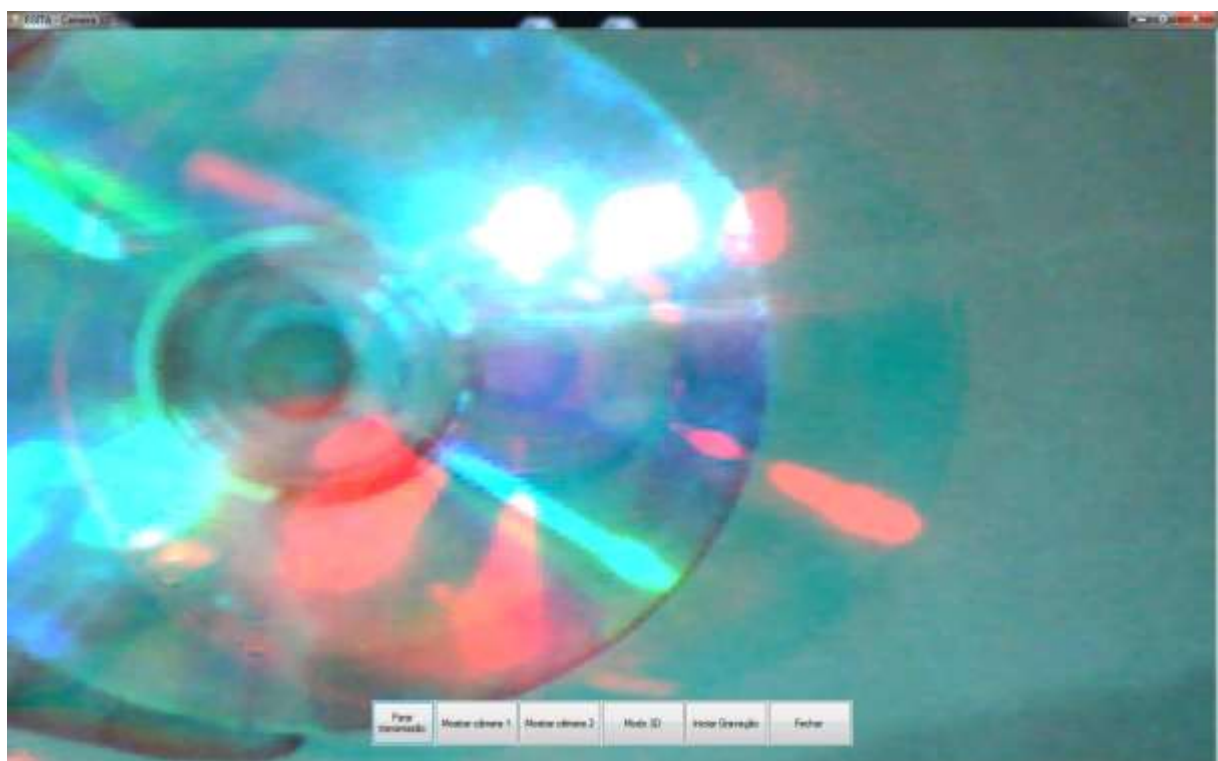
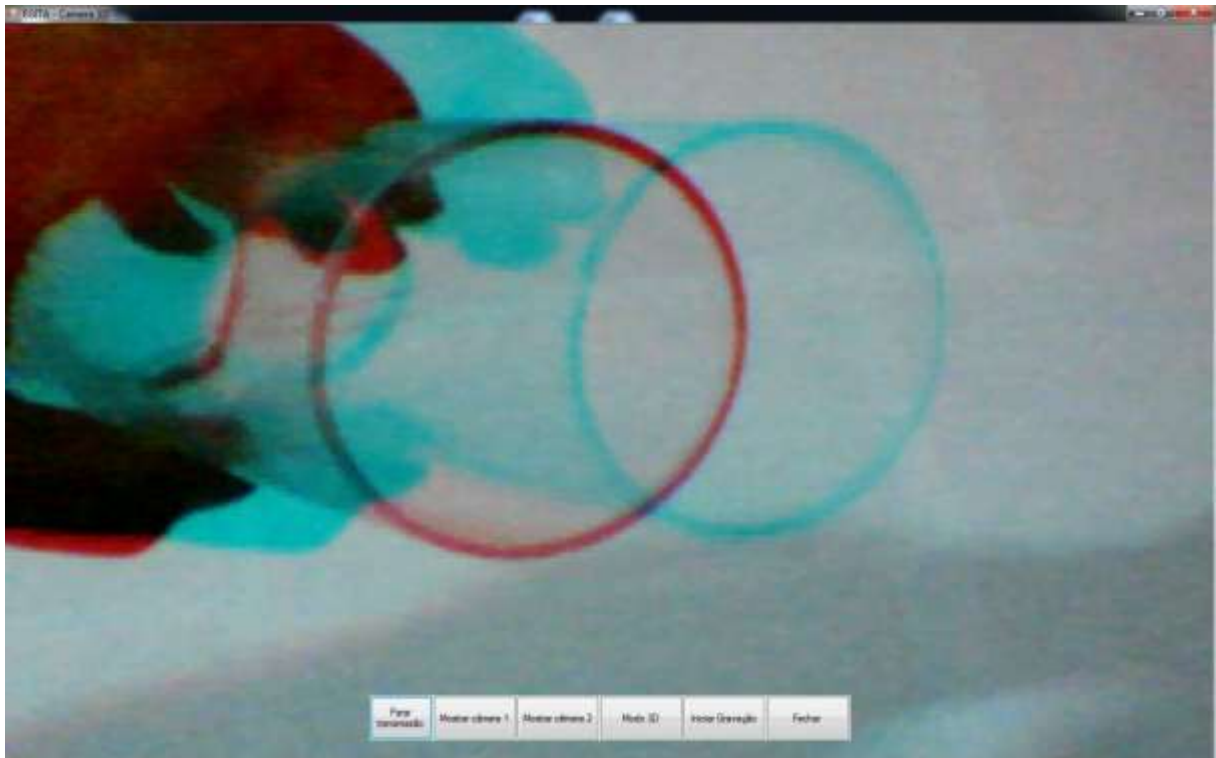
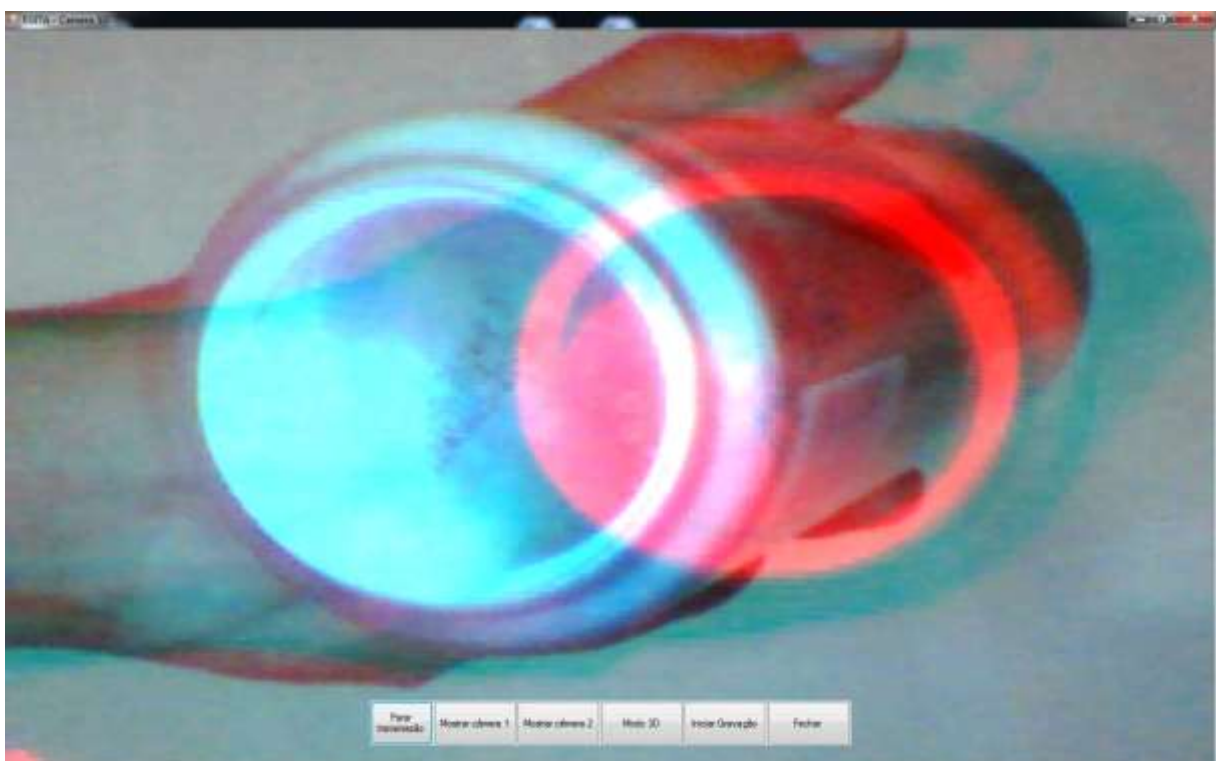


Figura 24 - CD com reflexo da luz 24cm





**Figura 25 - Copo de vidro 25cm**



**Figura 26 - Embalagem 24cm**

Com exceção da imagem “Figura 24 - CD com reflexo da luz 24cm” foi possível perceber profundidade na cena observada, também foi percebido que em objetos mais distantes não se nota a profundidade, mas ao observar a mesma imagem sem o óculos e voltando a observada novamente com o óculos, o objeto observado aparenta melhor realismo.

A imagem “Figura 26 - Embalagem 24cm” se destaca entre as outras, ao observar esta com o óculos 3D, é percebido parte do objeto saindo da tela com maior intensidade.

## 4. CONSIDERAÇÕES FINAIS

Ao iniciar o estudo sobre tecnologias 3D logo foi percebido o motivo pela qual se tanto ouve falar, os resultados obtidos fazem com que o estudante explore ainda mais a natureza dos assuntos que envolvem as tecnologias.

Dificuldades surgiram no desenvolvimento do sistema computacional, mesmo sendo conhecida a tecnologia 3D de interesse, no caso tecnologia 3D anágrifo. Capturar imagens da webcam não é algo simples, mesmo usando uma IDE contendo uma infinidades de recursos como o Visual Studio 2010. Como solução foi utilizada a biblioteca Emgucv para essa finalidade.

Os conhecimentos obtidos durante a formação influenciaram diretamente no planejamento e construção do sistema computacional, onde o objetivo foi alcançado e superado, pois as funções disponibilizadas na biblioteca Emgucv possibilitou a criação de um sistema computacional que aplica a tecnologia anáglifo em tempo de execução, ou seja, na medida em que as imagens das webcams são captadas, o sistema computacional processa estas imagens sem a preocupação de como foram adquiridas.

Com relação aos resultados do experimento, a eficácia da técnica estereoscópica anágrifo foi validada conforme os estudos, mas foi percebida uma dificuldade de percepção 3D na medida em que o tempo de observação aumentou, ou seja, ao observar imagens anaglíficas por algum tempo foi preciso mais e mais concentração para a percepção de profundidade.

## 5. REFERENCIAS

ANTÔNIO, Marco; CARLOS, José. **Fundamentos da visão computacional**. Presidente Prudente: Gráfica Viena, 2011.

AZEVEDO, David. **O sucesso do 3D em relançamentos de filmes**. Minas Gerais  
Disponível em: <http://www.intercom.org.br/papers/regionais/sudeste2012/resumos/R33-1469-1.pdf>. Acesso em: 22 mai de 2012

CASTRO, Clarissa. **Manipulação de imagens**.  
Disponível em: <http://vigo.ime.unicamp.br/~asaa/Clarissa.pdf>. Acesso em: 23 mai de 2012

GONZALEZ, Rafael; WOODS, Richard. **Processamento de imagens digitais** / Rafael C. Gonzalez, Richard E. Woods; tradução de Roberto Marcondes Cesar Junior, Luciano da Fontoura Costa. São Paulo: Blucher, 2000.

HOFFMAN, Donald. **Inteligência visual: Como criamos o que vemos** / Donald D. Hoffman; tradução de Denise Cabral de Oliveira. Rio de Janeiro: Campos, 2000.

LANDIM, Wikerson. **Como funciona a tecnologia 3D?**  
Disponível em: <http://www.tecmundo.com.br/video/2469-como-funciona-a-tecnologia-3d-.htm> Acesso em: 2 mai de 2012.

MORIN, Edgar. **Introdução ao pensamento complexo** / Edgar Morin; tradução de Eliane Lisboa. Porto Alegre: Suliana, 2006.  
Disponível em:  
<http://conhecaoliveiro.com.br/pdf/livros/cultura/833809.pdf> Acesso em: 22 abri de 2012

RAPOSO, Alberto, et al. **Visão Estereoscópica, Realidade Virtual, Realidade Aumentada e Colaboração**. Disponível em:  
[www.tecgraf.org.br/publications/artigo\\_2004\\_visao\\_estereoscopica\\_realidade\\_virtual.pdf](http://www.tecgraf.org.br/publications/artigo_2004_visao_estereoscopica_realidade_virtual.pdf).  
Acesso em: 26 jun de 2012.

SOUSA, Murilo, et al. **Estereoscopia aplicada à neuroanatomia**.  
Disponível em: <http://www.scielo.br/pdf/%0D/anp/v60n3b/a17v603b.pdf>. Acesso em: 27 jun de 2012.

WAN, Jorge; ANJOS, Daniel; MONTEIRO, Pedro. **Tecnologia 3D**. Portugal: Universidade Técnica de Lisboa, 2010.  
Disponível em:  
[http://www.img.lx.it.pt/~fp/cav/ano2010\\_2011/Trabalhos\\_MEEC\\_2011/Artigo%2016/CAV/resources/Artigo-de-Divulgacao.pdf](http://www.img.lx.it.pt/~fp/cav/ano2010_2011/Trabalhos_MEEC_2011/Artigo%2016/CAV/resources/Artigo-de-Divulgacao.pdf) Acesso em: 5 mai de 2012