

# Obtenção da disparidade e dos mapas de profundidade em vídeos 3D

Nathália Alves Rocha Batista <sup>[1]</sup>, Carlos Danilo Miranda Regis <sup>[2]</sup>

[1] nathaliaarb@gmail.com ; [2] regis.danilo@gmail.com. Instituto Federal de Educação, Ciência e Tecnologia da Paraíba - Campus Campina Grande.

## RESUMO

O Vídeo 3D é uma tecnologia baseada nas propriedades de percepção de profundidade humana. Duas perspectivas diferentes da mesma imagem são projetadas simultaneamente para cada um dos olhos do usuário, para fornecer a sensação de profundidade. A tecnologia 3D esta em crescente demanda e mais acessível devido à facilidade de aquisição, aos avanços tecnológicos em técnicas de visualização e transmissão, e de compressão de vídeo. Portanto, há a necessidade de avaliar os processos de codificação da profundidade dos vídeos 3D, para garantir a qualidade e a satisfação por parte dos usuários. Este artigo propõe uma análise sobre a forma de obtenção da profundidade a partir do cálculo da disparidade entre os vídeos. Os valores armazenados no Mapa de Disparidade representam a distância de cada pixel da imagem esquerda e direita, gerando-se um mapa de profundidade segundo uma escala de valores de cinza de 8 bits. De acordo com esses valores, é possível identificar a distância (profundidade) dos objetos da cena.

**Palavras-chave:** Vídeo 3D, Mapa de Disparidade, Profundidade, Estereoscopia, Foco.

## ABSTRACT

*3D video is a technology based on the properties of human depth perception. Two different perspectives of the same image are projected simultaneously to each of the user's eyes to provide a sense of depth. They are in increasing demand and more easily accessible due of acquisition, technological advances in visualization techniques and transmission, and video compression. Therefore, there is a need to evaluate the encoding processes of the depth of the 3D video, to ensure quality and greater satisfaction among users. This article proposes an analysis how to obtain the depth from the calculation of the disparity between the videos. The values stored in the Map of Disparities represents the distance of each pixel of the image left and right, generating a depth map using a scale of gray values of 8 bits. According to these values, one can identify the distance (depth) of scene objects.*

**Keywords:** 3D video, Disparity map, Depth, Stereoscopy, Focus.

## 1 Introdução

A estereoscopia é um fenômeno natural do Sistema de Visão Humano (HVS - *Human Visual System*) para a percepção da profundidade por meio da utilização de duas perspectivas diferentes de uma mesma imagem. Os displays 3D utilizam essa característica da visão humana para formação da imagem em terceira dimensão (CHAN *et al.*, 2007).

Os seres humanos utilizam características oculomotoras<sup>1</sup> para a percepção da imagem 3D. Dentre essas características destaca-se a acomodação, convergência, estereoscopia binocular, oclusão linear, tamanho relativo, borramento da retina, constroem a percepção de profundidade em uma cena 3D.

O processo de geração de vídeos em 3D simula o efeito de estereoscopia por meio de duas câmeras organizadas com a mesma distância inter-ocular. Logo, colocando-se as câmeras separadas uma da outra com base nessa distância, é possível simular o sistema visual humano. Quando cada imagem das câmeras for apresentada ao seu olho correspondente, as duas imagens serão fundidas em uma única imagem pelo cérebro, produzindo o efeito estereoscópico (JIAO e CHEOLKON, 2011).

Em termos matemáticos, a profundidade pode ser calculada a partir da disparidade entre os *pixels* da imagem direita e esquerda (FENG *et al.*, 2011). Profundidade e Disparidade estão relacionadas entre si, sendo uma inversamente proporcional à outra. Esse artigo se propõe a estudar os valores da disparidade gerados para alguns vídeos 3D.

Neste artigo, serão analisados três métodos de sensibilidade do HVS em relação à formação e percepção da profundidade em imagens 3D. Das características humanas na percepção da profundidade, uma relação entre disparidade e profundidade é demonstrada, bem como alguns exemplos do mapa de disparidade e mapas de profundidade das imagens.

As seções do artigo serão organizadas da seguinte forma: na Seção 2 tem a introdução sobre formação, visualização e codificação do vídeo 3D, apresentado um modelo matemático para o cálculo da disparidade entre duas imagens e a profundidade

de uma imagem; os resultados obtidos serão mostrados na Seção 3; e a conclusão na Seção 4.

## 2 Vídeo 3D

Na formação da imagem 3D são necessárias duas imagens, de uma mesma cena, referente à cada olho. As imagens são fundidas em um processo de codificação, gerando a imagem estereoscópica. As diferenças entre as imagens esquerda e direita (disparidade horizontal) referem-se à informação da distância que os objetos são percebidos na imagem, gerando diferentes profundidades.

A visualização das imagens 3D se dá pelo uso de óculos ou não, ou seja, as imagens podem ser multiplexadas temporalmente ou espacialmente. Os óculos utilizados podem ser do tipo anaglífico (utiliza filtro de cores) ou de luz polarizada (passivo ou ativo). Na tecnologia sem óculos, usa-se *display* autoestereoscópicos, nos quais são utilizadas películas especiais na superfície do *display*, de forma que a imagem é fatiada, residindo sobre as colunas pares e ímpares do *display* (COTILHO *et al.*, 2009).

Para obter um bom resultado na geração de vídeos em terceira dimensão, é necessário que haja um processamento otimizado das partes fundamentais como geração do conteúdo, codificação e transmissão, decodificação e exibição, e da percepção humana em 3D. A percepção 3D pode ser comprometida à medida que esses passos não são obtidos com êxito, causando um efeito conhecido com *crosstalk*. Esse problema pode ocorrer conforme a posição do telespectador, a qualidade da imagem, a tecnologia de processamento, ao ângulo de visão etc. O *crosstalk* causa um desconforto visual aos telespectadores (TSIRLIN *et al.*, 2011).

Vários parâmetros devem ser levados em consideração no momento em que se produz e codifica uma imagem 3D. A configuração das câmeras para captura das imagens é uma delas. As câmeras podem ser configuradas com eixo paralelo ou eixo convergente (*toded-in*). Na tecnologia 3DTV (COTILHO *et al.* 2009), a configuração paralela é mais utilizada, produzindo paralaxe horizontal da imagem. A configuração em *tode-in* pode gerar alguns erros na paralaxe vertical, ocasionando o efeito de *keystone*<sup>2</sup>.

1 Características responsáveis pela maioria da movimentação ocular extrínseca. Controla a movimentação dos músculos reto medial, reto superior, reto inferior, elevador da pálpebra superior e oblíquo inferior.

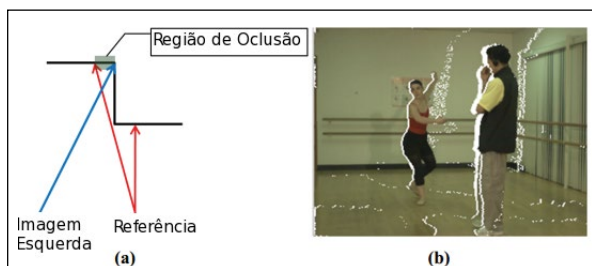
2 Imagem de forma trapezoidal em direções opostas para o olho esquerdo e direito na gravação da câmera. Em imagens estereoscópicas essas formas opostas da imagem trapezoidal induz paralaxe vertical e horizontal incorretas.

A aceitação do conteúdo 3D por parte do mercado consumidor não só depende dos avanços tecnológicos relacionados aos vários *displays* e suas tecnologias disponíveis no mercado, mas também sobre a disponibilidade de uma grande variedade de conteúdos 3D, conteúdos estes que podem ser obtidos a partir da captura de uma cena, com duas ou mais câmeras, ou a partir da conversão do conteúdo 2D em 3D (SILVA *et al.*, 2010, p 4013–4016).

Uma técnica utilizada atualmente para formação da imagem 3D é o DIBR (*Depth Image Based Rendering*), no qual a imagem 3D é formada a partir da multiplexação das imagens esquerda e direita resultantes da imagem de referência e o mapa de profundidade. Um problema inerente ao DIBR são as áreas de *desocclusão*<sup>3</sup>, ou seja, os pontos que são oclusos na imagem de referência tornam-se visíveis em alguns locais da imagem virtual esquerda e direita (JIAO e CHEOLKON, 2011).

Esse problema é solucionado na fase de pré-processamento da imagem, utilizando-se filtros passa-baixa (gaussianos). Esses filtros são capazes de reduzir de forma eficiente as áreas de *desocclusão* e manter um padrão de qualidade na criação de imagens estereoscópicas.

**Figura 1** – Os “buracos” causados pela desocclusão. (a) Causa da desocclusão (b) Imagem esquerda de “Ballet” com holes



## 2.1 Relação da profundidade e sensibilidade da visão humana

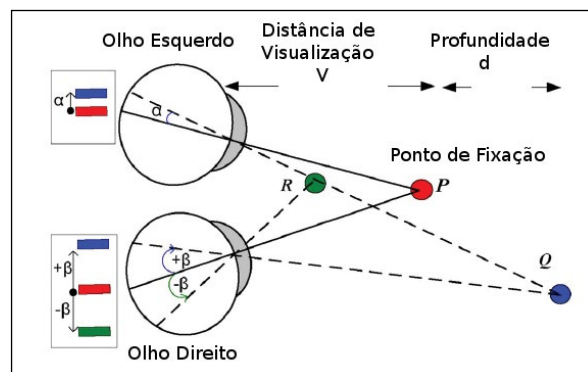
Os tópicos a seguir apresentam três formas com as quais o cérebro humano interpreta a imagem 3D e percebe a profundidade da imagem.

### 2.1.1 Estereoscopia binocular

A estereoscopia ocorre devido ao fato dos olhos estarem horizontalmente separados, aproximadamente 6,3 cm, proporcionando a cada olho, um ponto de vista único sobre uma imagem (ZHANG; TAM, 2005). A imagem é formada na região central da retina (fóvea) de tal forma que cada olho observa a imagem em ângulos diferentes. A diferença entre os ângulos é chamada de disparidade binocular e fornece informações sobre a distância relativa dos objetos até o observador, a estrutura de profundidade e o ambiente em geral (SILVA *et al.*, 2010).

Na Figura 2 é demonstrado como a imagem é formada na retina do telespectador.

**Figura 2** – Geometria da Estereoscopia Binocular



Fixando-se um ponto P no centro de cada olho, fixamos outros dois pontos, Q e R, de forma que Q esteja localizado depois do ponto central e R antes de ponto central. Cada imagem projetada por esses pontos formará ângulos em relação ao ponto central, caracterizando a disparidade angular em que Q tem disparidade positiva e R negativa. Essa diferença de sinal estimula ao cérebro a perceber a profundidade dos objetos em relação à posição deles (SILVA *et al.*, 2010).

A disparidade angular de R e Q pode ser obtida por:

$$\eta_Q = (\beta - \alpha). \tag{1}$$

Similarmente, podemos calcular a disparidade angular no ponto R da seguinte forma:

3 Problema relacionado à visualização de pontos sem textura nas imagens virtual esquerda e direita.

$$\eta_R = -(\beta) - \alpha = -(\beta + \alpha). \quad (2)$$

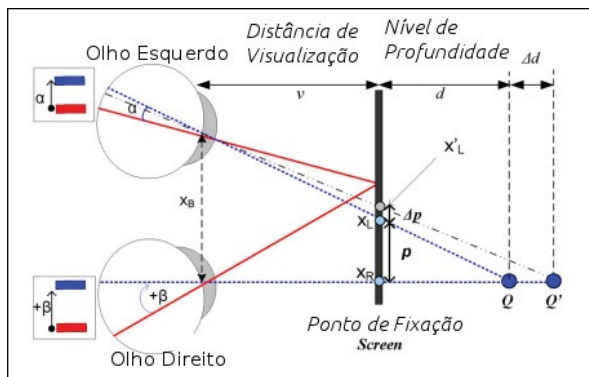
Analisando geometricamente a Figura 2, observa-se que a disparidade é diretamente proporcional à profundidade  $d$  e inversamente proporcional ao quadrado da distância  $v$ .

$$\eta \propto \frac{d}{v^2} \quad (3)$$

De acordo com a Equação 3,  $\eta_Q > 0$  e  $\eta_R < 0$ , cérebro interpreta esta diferença de sinais como o posicionamento relativo dos pontos Q e R em relação ao ponto P. Quando a disparidade binocular é maior do que zero, a imagem é formada atrás do ponto de fixação. Já para a disparidade menor que zero, a imagem é formada na frente do ponto de fixação. Assim, a disparidade binocular fornece estímulo ao cérebro para perceber a profundidade de objetos em relação ao ponto de fixação.

A Figura 3 relaciona a sensibilidade da visão humana com a formação da imagem em um display estereoscópico.

Figura 3 – Estereoscopia Binocular e Display 3D



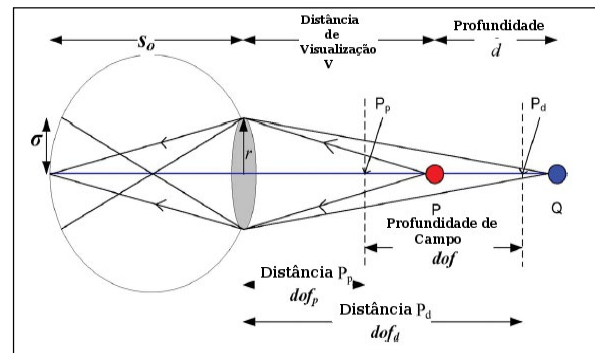
### 2.1.2 Borrimento da retina e campo de profundidade

O borrimento da retina é um método visual no qual a profundidade é percebida com apenas um olho. O borrimento de uma imagem simula a percepção da profundidade, pois os olhos automaticamente

focam (acomodação) no objeto fixado, fazendo com que este objeto sobressaia dentre os outros objetos ao seu redor, provocando um borrimento nos objetos mais distantes (SILVA *et al.*, 2010).

Na Figura 4 pode ser observada a forma geométrica para análise da formação da imagem na retina, por meio do borrimento da imagem.

Figura 4 – Geometria do Desfoque na Retina



Fixando o ponto P, a capacidade óptica (convergir ou divergir a luz) dos olhos é ajustada para que a imagem do ponto P esteja focada na região central da retina (fóvea). No entanto, a capacidade óptica dos olhos varia em certa frequência e, assim, o olho pode tolerar certa quantidade de desfoque da retina sem reajuste de acomodação.

Esta diferença de capacidade óptica é conhecida como a profundidade de foco ocular e é apresentada em dioptrias<sup>4</sup> (D). Em outras palavras, o ponto P pode ser deslocado ao longo do eixo óptico dentro de um período determinado, sem perceber o desfoque da imagem. Este período é conhecido como a profundidade de campo *dof*.

Os pontos mais próximos e mais distantes do limite são conhecidos como ponto "Proximal" ( $P_p$ ) e ponto "Distal" ( $P_d$ ), como mostrado na Figura 3. Portanto, qualquer objeto dentro do limite é percebido nitidamente e os objetos fora do limite são percebidos "borrados", o que estimula a sensação de Profundidade (SILVA *et al.*, 2010, p 4013-4016). A Equação 4 relaciona a distância do olho ao ponto  $Q(v+d)$  e o valor do borrão da retina  $\sigma$ :

$$Q(v+d) = \frac{F.r.S_o}{r.S_o - F(r+\sigma)} \quad (4)$$

4 Dioptria é uma unidade de medida que afere a capacidade de divergência de um sistema óptico ( $m^{-1}$ ). Na Óptica, é a unidade de medida da potência de uma lente corretiva (popularmente conhecido como grau).

Em que  $v$  representa a distância visual,  $d$  a profundidade,  $F$  o tamanho focal das lentes,  $r$  o raio de abertura dos olhos e  $S_o$  a distância do centro da retina (fóvea) aos olhos.

A magnitude do campo de profundidade ( $dof$ ) difere de pessoa para pessoa, dependendo da profundidade do foco ocular. A Equação 5 representa o cálculo da profundidade do foco ocular  $T$  relacionando o ponto "Proximal" e "Distal", de acordo com os padrões das equações ópticas,

$$T = \frac{1}{dof_p} - \frac{1}{dof_d}, \quad (5)$$

e os valores de  $dof_p$  e  $dof_d$  são dados por:

$$dof_p = \frac{2v}{(2 + v.T)} \quad (6)$$

$$dof_d = \frac{2v}{(2 - v.T)} \quad (7)$$

Não há nenhuma indicação na literatura sobre como calcular os valores de  $T$ . No entanto, ele varia entre 0.6 D e 0.8 D. Portanto, quando um objeto é nítido e o outro é borrado tem-se a sensação além da distância do objeto nítido. Desta forma, são definidos limites para os objetos projetados em uma tela estereoscópica para minimizar o desconforto visual.

### 2. 1.3 Tamanho relativo

O tamanho relativo da imagem é outro método visual no qual a profundidade é percebida pela imagem formada na retina. O ângulo visual de um objeto projetado na retina diminui à medida que a distância até o objeto aumenta e vice-versa. Utilizando a propriedade de semelhança dos triângulos, o tamanho da imagem na retina  $R$  em função do tamanho real da imagem  $H$ , distância focal  $G$  e da distância do objeto  $D$  se relacionam como demonstra a Equação 8:

$$\frac{R}{F} = \frac{H}{D} \Rightarrow R = \frac{H}{D} \cdot F, \quad (8)$$

onde os valores de  $D$  referem-se à distância física.

O aumento na distância do objeto interfere no tamanho da imagem formada na retina, diminuindo-a de tamanho. O cérebro interpreta essas diferenças

como uma mudança de profundidade  $\Delta R$ . A Equação 9 representa matematicamente esse comportamento:

$$R - \Delta R = \frac{H}{(D + \Delta d)}. \quad (9)$$

É possível estimular uma mudança no tamanho da imagem projetada na retina por meio da mudança do tamanho do objeto  $\Delta H$ . Com isso, as mudanças no tamanho dos objetos que são fisicamente similares proporcionam a sensação de profundidade devido a uma mudança na retina do tamanho da imagem.

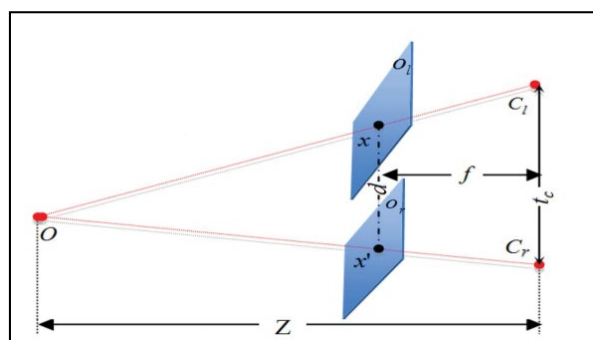
$$R - \Delta R = (H - \Delta H) \frac{F}{D} \quad (10)$$

## 2.2 Disparidade X Profundidade

Disparidade refere-se às diferenças espaciais entre duas imagens, esquerda e direita, e é definida como sendo a *distância horizontal* entre dois *pixels* correspondentes. Eles usualmente se encontram em posições diferentes nos *frames* ou pode acontecer que o *pixel* não tenha seu homólogo. Determinar o valor da disparidade é essencial para se determinar a distância dos objetos às câmeras.

Dada a informação da disparidade associada a qualquer par de imagens, o cérebro pode gerar a percepção de profundidade, fundindo essas imagens. A ideia de gerar a percepção de profundidade a partir de informações da disparidade inspirou a captura de imagens estereoscópicas, de uma mesma cena, usando duas câmeras com a mesma configuração. A relação entre profundidade e disparidade em uma câmera para captura das imagens, é mostrada na Figura 5.

Figura 5 – Geometria da Câmera





Os dados do cálculo da disparidade são armazenados na forma de mapas de disparidades, que são vetores de informação com os valores da distância do *pixel* horizontal para cada coordenada de *pixel* da imagem. Podem ser representados também por uma imagem resultante de todas as diferenças entre os *pixels*, que são os mapas de profundidade. Alguns problemas como oclusão podem dificultar a obtenção da disparidade (CHAN *et al.*, 2007).

O valor da disparidade entre todos os *pixels* da imagem direita e esquerda é obtido pela Equação 11. Nessa equação, são calculadas as diferenças entre dois *pixels* homólogos, chamando de *Método de Correspondência*.

$$d = |X_l - X_r| \quad (11)$$

O *Método de Correspondência* consiste em determinar um *pixel* homólogo da imagem esquerda, na imagem direita. A determinação desses *pixels* homólogos depende da qualidade da imagem e da precisão com que a cena foi reconstruída. Os mapas de profundidade podem ser classificados como mapas esparsos ou mapas densos. Os mapas esparsos são obtidos a partir do cálculo da disparidade para alguns pontos e os mapas densos são obtidos pelo cálculo da disparidade por todas as áreas da imagem.

Uma vez determinado o valor da disparidade entre os *pixels* da imagem, pode-se converter esse dado em distâncias físicas (CHAN *et al.*, 2007). Para se determinar a profundidade *Z* com base nos valores de disparidade, é necessário conhecer os parâmetros físicos das câmeras, *B* e *f*, como demonstrado na Equação 12:

$$Z = \frac{B \cdot f}{d} \quad (12)$$

, onde *B* representa a distância entre as câmeras e *f* representa a distância focal de cada câmera. O valor de *B* e *f* são sempre positivo e, portanto para profundidades positiva ou infinita, a disparidade *d* é sempre positiva ou nula. Os valores de *B* e *f* devem ser descritos nas especificações de cada vídeo, pois são valores inerentes a eles levando em consideração a estrutura de filmagem (ZHANG *et al.*, 2005).

Convém ressaltar que o valor da disparidade pode ser positivo, negativo ou nulo. No caso de uma disparidade nula, entende-se que foi encontrada uma distância tal que já não se consegue distinguir uma

diferença de posição, pois aquele ponto encontra-se na mesma posição para ambas as câmeras. Na disparidade positiva, a imagem será formada "para dentro" do plano do *display*, enquanto que a disparidade negativa será formada "para fora" do plano do *display*.

### 3 Resultados e discussão

Nesse trabalho, verificamos os valores da disparidade para alguns vídeos 3D disponibilizados por instituições de pesquisa na área de processamento de imagens. O método para estimativa da profundidade utilizado neste trabalho foi o *depth-from-disparity*, ou seja, a informação de profundidade é obtida através do cálculo da disparidade entre as vistas esquerda e direita.

Para obtenção dos resultados, foram utilizados vídeos em *.YUV*, com resolução de 432x240 e *.AVI* com resolução 1856x1036. Os vídeos em *.AVI* foram decodificados para *.YUV* utilizando os softwares *YUV Tools*, *Mencoder* e o *JM* versão 18.4.

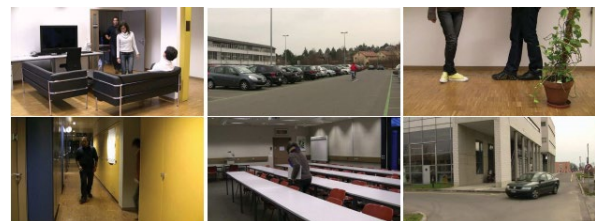
Na Figura 6 são apresentadas as imagens de formato *yuv:4:2:0*. Esse banco de dados foi disponibilizado por *Mobile 3DTV* (SMOLIC *et al.*, 2011).

Figura 6 – Banco de Dados 1



O banco de dados 2 foi disponibilizado pela *Ecole Polytechnique Federale de Lausanne (EPFL)* (GOLDMANN *et al.*, 2010).

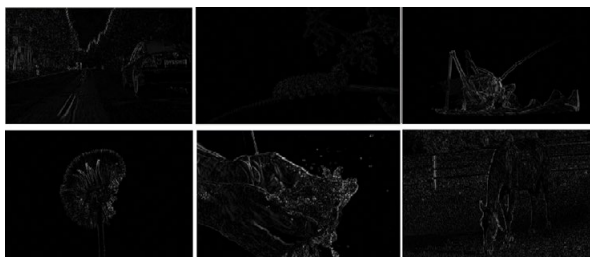
Figura 7 – Banco de Dados 2



Os mapas de profundidades foram obtidos a partir do cálculo da disparidade entre as imagens

esquerda e direita. Um algoritmo em C foi implementado, no qual são utilizados como parâmetro os vídeos referentes ao olho esquerdo e direito, gerando em arquivo de texto todos os valores da disparidade entre os *pixels*.

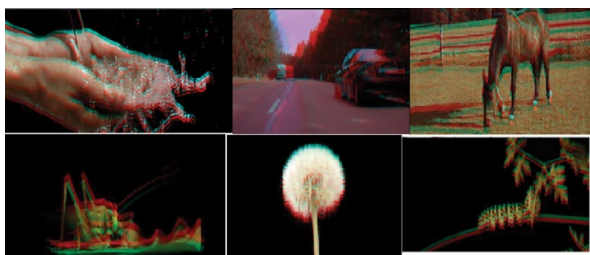
**Figura 8** – Mapas de Profundidade das Imagens do Banco de Dados .YUV



Dos valores obtidos, foi observado que a disparidade varia dentro do intervalo  $[-128... 127]$ , e que os mapas de profundidade são imagens de 8 *bits* na escala cinza, no qual as cores mais claras têm valores de disparidades maiores e as cores mais escura, valores de disparidade menores (Figura 8).

Para visualização 3D utilizando óculos anaglíficos, os vídeos foram decodificados usando software JM Reference versão 18.4.

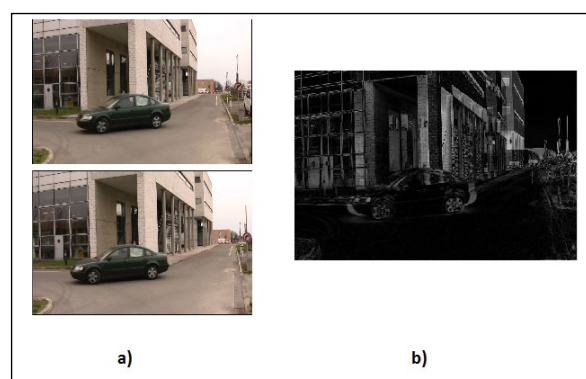
**Figura 9** – Vídeo codificados para formar a imagem 3D anaglífica



Dos valores de disparidades calculados para os vídeos do banco de dados 1, no vídeo "Hands.yuv", foram encontrados valores positivos, negativos e nulos, de forma que os três tipos de profundidade puderam ser observados. O vídeo "Car.yuv" e "Horse.yuv" tiveram valores de disparidade predominantemente positivos, indicando que a profundidade desse vídeo é "para dentro" do display. Já os vídeos "Grasshopper.yuv", "Flower.yuv" e "Caterpillar.yuv", tiveram valores de disparidade positivos e iguais a zero, indicando que as profundidades são projetadas no plano do display e "para dentro" do display.

Nas imagens disponíveis no banco de dados 2, o vídeo "Car-outdoor", de acordo com os resultados do mapa de disparidade, esse vídeo apresenta os três tipos de profundidade, pois os valores das disparidades resultantes são negativos, positivos e zero. Apesar de serem encontrados valores negativos, a profundidade predominante nessa imagem é "para dentro" do display (Figura 10).

**Figura 10** – (a) Imagens Esquerda e Direita (b) Mapa de Profundidade



Profundidade em vídeos 3D é um assunto amplamente estudado ao longo dos últimos anos e vários métodos para extração da profundidade foram estudados e documentados. Métodos para extração da profundidade disponíveis na literatura são baseados em geometria, movimento, segmentação, classificação cena e Inteligência Artificial. Foi utilizado nesse trabalho uma análise geométrica da profundidade, pois a profundidade é inversamente proporcional ao valor da disparidade e os parâmetros do tamanho focal  $f$  e a distância entre as câmeras  $B$  são facilmente acessíveis. Além disso, a escolha das imagens para estimativa da profundidade nesse método é muito importante, pois em imagens pouco esparsas a profundidade é melhor observada para objetos mais próximos e, para imagens mais esparsas a profundidade é melhor percebida para objetos mais distantes. As imagens escolhidas apresentam profundidades de todos os tipos, de acordo com o cálculo da disparidade, o que proporcionou uma visão geral sobre os tipos de profundidade existentes.

## 4 Conclusão

Neste artigo, apresentamos uma análise da sensibilidade humana na percepção da profundidade que proporciona a visualização em displays 3-D.

Na estereoscopia binocular, o telespectador percebe a profundidade a partir da diferença das imagens formadas nas duas retinas. Na sensibilidade, a imagem borrada na retina, o usuário percebe a profundidade da imagem de acordo com o borramento de objetos mais distantes do plano da imagem, de forma que os objetos mais próximos fiquem mais visíveis. Já no tamanho relativo do objeto, a profundidade é vista pela diferença do tamanho dos objetos. O tamanho dos objetos pode ser aumentado ou diminuído à medida que os objetos ficam mais próximos ou distantes do plano da imagem. Em contrapartida, se o objeto tiver uma redução brusca de tamanho, podem ser observadas nas imagens áreas de desocclusão. Portanto, essas características são fundamentais na percepção da profundidade e na naturalidade dos conteúdos produzidos.

Utilizando características relativas à estereoscopia binocular, é apresentada uma forma para calcular a disparidade das imagens esquerda e direita para obter o mapa de profundidade da imagem 3D. Existem alguns métodos mais robustos para calcular os mapas de profundidade como, por exemplo, o software DERS (*Depth Estimation Software*), que fornece uma estimativa da profundidade e a síntese das vistas direita e esquerda, como é realizado na exploração de experimentos (EE) em 3D Vídeo(3DV) com o MPEG-FTV(Free-viewpoint TV). Esse algoritmo foi desenvolvido pela Nagoya University Tanimoto, utilizando algoritmos de *Stereo Matching* e *Graph Cuts*.

Levando em consideração a qualidade da imagem codificada, alguns métodos de filtragem podem reduzir a quantidade de ruídos nas imagens de referência, melhorando os resultados obtidos. Na literatura, alguns autores utilizam filtros assimétricos e filtros passa-baixa (Gaussiano). Neste trabalho não foi utilizado nenhum método de redução de ruídos que ocasionam os problemas de *holes*<sup>5</sup> nas imagens (SHAO *et al.*, 2009, pg 1-4).

Em trabalhos futuros pretendemos utilizar os outros métodos para estimar a profundidade, avaliar a profundidade da imagem a partir dos valores obtidos dos mapas de disparidades, gerar valores

para quantizá-la, definindo quais tipos é preferível pelos usuários, levando em consideração a avaliação tanto subjetiva quanto objetiva. Além disso, serão utilizados filtros na fase de pré-processamento das imagens, para que os problemas em decorrência da desocclusão sejam reduzidos.

## REFERÊNCIAS

- CHAN, S.C.; SHUM, H.Y.; NG, K. Image-based rendering and synthesis. *IEEE Signal Processing Magazine*, 24(6):22–33, November 2007.
- COTILHO, R. F.; SILVA, A. S.; CARINHAS, T. L.. *Televisão 3D, Comunicação Áudio e Vídeo*, Instituto Superior Técnico, Lisboa – Portugal, 2009.
- FENG, Y.; REN, J.; JIANG, JIANMIN, Object-Based 2D-to-3D Video Conversion for Effective Stereoscopic Content Generation in 3D-TV Applications, *IEEE Transactions on broadcasting*, vol. 57, no. 2, June 2011.
- GOLDMANN, L.; SIMONE, D.F.; EBRAHIMI, T. A Comprehensive Database and Subjective Evaluation Methodology for Quality of Experience in Stereoscopic Video, *Ecole Polytechnique Federale de Lausanne (EPFL)*, CH-1015 Lausanne, Switzerland.
- JIAO, L.C.; CHEOLKON, J. Disparity-map-based rendering for mobile 3D TVs, *IEEE Transactions on Multimedia*, September 2011.
- SHAO, H.; CAO, X.; ER, G. Objective quality assessment of depth image based rendering in 3dtv system. In *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video*, 2009, pages 1–4, May 2009.
- SILVA, D.V.S.X.; FERNANDO, W.A.C.; G.; NUR; EKMEKCIOGLU, E.; WORRAL, S.T. 3D video assessment with just noticeable difference in depth evaluation. In *17th IEEE International Conference on Image Processing (ICIP)*, pages 4013–4016, September 2010.
- SMOLIC A.; WANG, Y.; MÜLLER, M.; KAUFF P. Report on generation of video plus depth data base, *MOBILE3DTV*, 2011.
- TAM, W.J.; SPERANZA, F. Y.; SHIMONO, S.K.; ONO, H. Stereoscopic 3d-tv: Visual comfort. *IEEE Transactions on Broadcasting*, 57(2):335–346, June 2011.

<sup>5</sup> Pontos sem texturas nas imagens de referência.



TSIRLIN, I.; WILCOX, L. M.; ALLISON, R. S. The Effect of Crosstalk on the Perceived Depth From Disparity and Monocular Occlusions, IEEE Transactions on Broadcasting, vol. 57, no. 2, June 2011.

ZHANG, L.; TAM, W.J. Stereoscopic image generation based on depth images for 3d TV. IEEE Transactions on Broadcasting, 51(2):191 – 199, June 2005.

ZHANG, L.; TAM, W.J.; WANG, D. Stereoscopic image generation based on depth images. In International Conference on Image Processing, 2004. ICIP '04, volume 5, pages 2993 – 2996 Vol. 5, October 2004.

ZONE, R. Stereoscopic cinema and the origins of 3-D film, 1838-1952. The University Press of Kentucky, Kentucky, USA, 2007.