

# Uma Análise da Influência da Subamostragem de Crominância em Vídeos Estereoscópicos Anaglíficos<sup>1</sup>

Leonardo Antonio de Andrade  
Universidade Federal de São Carlos,  
Departamento de Artes e Comunicação  
Rod. Washington Luiz, Km 235  
São Carlos, São Paulo, Brasil  
landrade@ufscar.br

Rudinei Goularte  
Universidade de São Paulo,  
Instituto de Ciências Matemáticas e de Computação  
Av. Trabalhador São Carlense, 400  
São Carlos, São Paulo, Brasil  
rudinei@icmc.usp.br

## ABSTRACT

This paper presents an analysis on how the chrominance subsampling affects the three-dimensional perception in stereoscopic anaglyph videos. The analysis used objective and subjective measures of video quality, encoded by color matching red/cyan and green/magenta methods. The chrominance subsampling 4:2:2 and 4:2:0 (at color space YCbCr) was applied for each method of color combination and the results have shown that the method green/magenta is less sensitive than red/cyan method for chrominance subsampling. The analysis also demonstrates that subsampling 4:2:0 is prohibitive for both methods.

## Categories and Subject Descriptors

I.2.10 [Vision and Scene Understanding]: 3D/stereo scene analysis.

## General Terms

Algorithms, Measurement, Experimentation, Verification.

## Keywords

Digital Video, Stereoscopia, Anaglyph Stereoscopia Vision, Chrominance Subsampling.

## RESUMO

Este trabalho apresenta uma análise sobre como a subamostragem de crominância afeta a correta percepção tridimensional em vídeos estereoscópicos anaglíficos. A análise utilizou medidas objetivas e subjetivas de qualidade sobre vídeos codificados pelos métodos de combinação de cores vermelho/ciano e verde/magenta. Para cada método de combinação de cores, foram aplicadas subamostragens de crominância 4:2:2 e 4:2:0, e os resultados mostraram que o método verde/magenta sofre menos interferência da subamostragem de crominância do que o método vermelho/ciano, e que a subamostragem 4:2:0 é proibitiva para os dois métodos.

## Palavras-Chave

Vídeo Digital, Estereoscopia, Visão Anaglífica Estereoscópica, Subamostragem de Crominância.

## 1. INTRODUÇÃO

A visão estereoscópica é baseada na capacidade de percepção da profundidade, pela aquisição de duas imagens da mesma cena horizontalmente separadas, representando a visão dos olhos esquerdo e direito [1]. O cérebro humano processa as diferenças entre as imagens adquiridas pelos olhos para perceber a profundidade na cena [2]. É designado par estéreo as duas imagens que simulam as imagens adquiridas pela visão de ambos os olhos [3].

O vídeo estereoscópico, também denominado de vídeo estéreo, é um tipo de vídeo composto por seqüências de pares estéreo. Para a percepção do vídeo estéreo é necessário um sistema especial de visualização, que na maioria das vezes utiliza óculos específicos (anaglíficos, polarizados ou obturadores), para realizar a separação das visões direita e esquerda sendo exibidas em monitor ou projeção [4]. A exceção são os monitores auto-estereoscópicos, no qual não são necessários óculos especiais [5][6].

Apesar do recente impulso que a tecnologia 3D vem recebendo da indústria do cinema [7] [8] e da televisão [9], ainda existe necessidade de mais pesquisa na área. Um reflexo disso é a atual falta de padronização no modo de organizar os dados estéreo para fins de armazenamento ou transmissão, sendo que tais métodos podem ser divididos em dois grandes grupos: os métodos aqui chamados de vinculados [10] e o método de Lipton [11].

No método de Lipton o par estéreo é armazenado em containers (AVI por exemplo), com compressão ou não. Apesar de ser mais flexível que os métodos vinculados, resulta em um volume de dados duas vezes maior, devido à necessidade de se armazenar dois streams de vídeo (o par estéreo).

Em direção ao estabelecimento de um método de codificação de vídeos estéreo que seja flexível, não vinculado e com baixo volume de dados, são necessárias ferramentas que atendam aos requisitos de todos os métodos de visualização estéreo. Nesse ponto existe uma problemática particularmente ligada ao método anaglífico.

A visualização anaglífica tem sido utilizada em diversas áreas, tais como visualização científica [12], visualização médica [13], ferramenta de ensino [14][15], e a na área de cinema [16], mostrando que ainda existe espaço para essa tecnologia que tem baixo custo comparada a outras técnicas [4]. Com este método é possível utilizar um único stream de vídeo, economizando assim

<sup>1</sup> An Analysis of Chrominance Subsampling Influence in Anaglyph Stereoscopic Videos

espaço de armazenamento e facilitando sua transmissão. No entanto, qualquer que seja o armazenamento do vídeo estéreo, ele necessita de compressão e neste procedimento nos deparamos com duas técnicas, com e sem perdas. A utilização da compressão sem perdas não é suficiente para a demanda atual das aplicações multimídia envolvendo vídeos [17], e isso se torna mais complexo com os vídeos estereoscópicos, devido à necessidade de maior espaço de armazenamento. A compressão com perdas gera diferenças nas cores, geralmente imperceptíveis em um vídeo comum. Porém, tais perdas podem impossibilitar a correta visualização de vídeos estereoscópicos anaglíficos [18].

Assim, o objetivo deste trabalho é analisar uma etapa comum da compressão de vídeos, a subamostragem de crominância, verificando como esta afeta a visualização anaglífica. Em particular a análise busca apontar qual o melhor método de combinação de cores (verde/magenta ou vermelho/ciano) e qual o método mais adequado para realizar a subamostragem (4:2:2 ou 4:2:0). Os resultados obtidos trazem contribuições para o desenvolvimento de técnicas mais efetivas de compressão de vídeos estereoscópicos anaglíficos.

Este texto está organizado do seguinte modo: a seção 2 mostra a estereoscopia anaglífica aplicada a vídeos digitais; a seção 3 discute os modelos de compressão de vídeo, e as etapas em que esse processo ocorre; a seção 4 mostra uma base de testes estereoscópica anaglífica e a escolha de técnicas de subamostragem de crominância utilizadas para testes de compressão; a seção 5 mostra as medidas objetivas e subjetivas de qualidade aplicada aos vídeos codificados; por último, a seção 6 discute as conclusões do trabalho.

## 2. ESTEREOSCOPIA ANAGLÍFICA

O método de visualização estereoscópica anaglífica é o mais simples dos métodos de visualização estereoscópica. Esta técnica caracteriza-se em colorizar com uma cor primária diferente cada uma das imagens referentes a cada olho, de modo que o espectador tenha sensação da profundidade. A primeira experiência tridimensional utilizando esse método data de 1922, com o filme “Power of Love” [3].

Esta técnica apresenta-se como uma alternativa viável para visualização 3D de vídeos, pois necessita apenas guardar informações em um vídeo monocular com três canais de cores (RGB).

Considerando o par de vídeos estéreo, temos suas componentes definidas como  $R_1G_1B_1$  para as componentes de cor do vídeo da visão do olho direito e  $R_2G_2B_2$  para as componentes de cor do vídeo da visão do olho esquerdo. O vídeo anaglífico combina as componentes de cor do par estéreo em um novo vídeo. As combinações possíveis de cor para este tipo de formato de armazenamento são dadas a seguir:

- Vermelho/Ciano: neste caso o vídeo estereoscópico anaglífico utiliza as componentes verde e azul do vídeo da visão do olho direito e a componente vermelha da visão do olho esquerdo ( $R_2G_1B_1$ );

- Verde/Magenta: neste caso o vídeo estereoscópico anaglífico utiliza as componentes vermelha e azul do vídeo da visão do olho direito e a componente verde da visão do olho esquerdo ( $R_1G_2B_1$ );

- Azul/Amarelo: neste caso o vídeo estereoscópico anaglífico utiliza as componentes verde e vermelha do vídeo da visão do

olho direito e a componente azul da visão do olho esquerdo ( $R_1G_1B_2$ ).

Neste escopo, para visualizar o vídeo estéreo anaglífico é necessário utilizar óculos com as respectivas lentes de filtros coloridos para separar cada uma das imagens que se encontram misturadas no monitor ou projeção. Os filtros de cor precisam ser complementares para uma precisa visualização do vídeo anaglífico [19]. Isso implica que o primeiro filtro deve deixar determinados comprimentos de onda passar e barrar os outros do espectro visível. O outro filtro tem ação complementar a essa, barrando as informações que o primeiro deixou passar e deixando passar as que o primeiro barrou. Se ambos os filtros deixaram determinado comprimento de onda passar, essa informação será passada aos dois olhos, e isso pode gerar imperfeições que dificultam a percepção estereoscópica.

Para os métodos que combinam as cores vermelho/ciano e verde/magenta, foram pesquisados filtros padronizados que correspondem aos filtros comumente utilizados para óculos encontrados no mercado [19]. Os códigos dos filtros Roscolux encontrados se encontram a seguir:

- Vermelho: #25 – Orange Red;
- Ciano: #69 – Brillant Blue;
- Verde: #139 – Primary Green.
- Magenta: #339 – Broadway Pink;

Pela dificuldade de serem encontrados no mercado filtros para o método amarelo/azul, não foi possível sua análise no presente trabalho.

A visualização anaglífica com o método verde/magenta, apresenta resultados melhores do que o método vermelho/ciano. Esse fato está relacionado à complementação dos filtros, como mostram as Figuras 1 e 2 nos gráficos de comprimento de onda do espectro eletromagnético visível (que resultam nas cores visíveis) pela porcentagem da transmissão de luz.

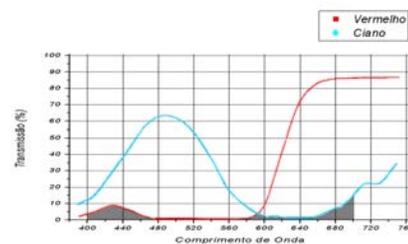


Figura 1: Sobreposições dos gráficos de comprimento de onda do espectro eletromagnético visível pela porcentagem da transmissão de luz das cores ciano e vermelho.

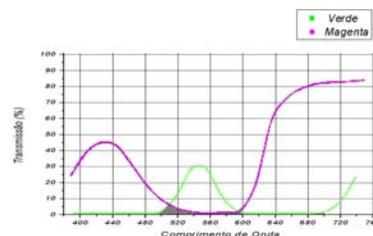


Figura 2: Sobreposições dos gráficos de comprimento de onda do espectro eletromagnético visível pela porcentagem da transmissão de luz das cores magenta e verde.

Podemos observar que nos dois gráficos anteriores, a área escura representa uma região onde os filtros não são complementares. Na Figura 1 a área escura é maior do que a área escura da Figura 2, mostrando que a combinação verde/magenta possui filtros complementares melhores do que da combinação vermelho/ciano.

### 3. TRABALHOS RELACIONADOS

Relacionado à percepção de vídeos estereoscópicos anaglíficos, existe o trabalho de Eric Dubois no qual foi desenvolvida uma metodologia para o método vermelho/ciano [19]. A solução proposta consiste em um processamento sobre os canais de cor RGB do vídeo, gerando assim componentes de cor filtradas. O trabalho desenvolvido no presente artigo não incluiu a possibilidade de pós-processamento dos canais de cor, realizando a visualização dos vídeos sem nenhum processamento adicional.

Outro trabalho relacionado a percepção estereoscópica em vídeos anaglíficos, foi realizada em 2009 por Andrade e Goularte [18], onde a percepção em vídeos anaglíficos vermelho/ciano foi analisada em vídeos com crominância e em vídeos apenas com luminância, onde os vídeos com crominância tiveram sua percepção estereoscópica alterada ao sofrerem codificações com perda.

O presente trabalho mostra uma comparação entre métodos de combinação de cor, analisando as possibilidades técnicas de ambas as combinações (vermelho/ciano e verde/magenta) para vídeos digitais. Além da análise de complementaridade dos filtros (Figura 1 e 2), será analisada a subamostragem da crominância de vídeos estereoscópicos anaglíficos. Ainda hoje, principalmente na área de ferramenta de ensino [14][15], a tecnologia anaglífica com a combinação vermelho/ciano é empregada. A comparação das técnicas pode trazer futuros benefícios para melhorar a qualidade da percepção de imagens e de vídeos digitais anaglíficos.

### 4. SUBAMOSTRAGEM DE CROMINÂNCIA

Uma etapa do processo de compressão de vídeos digitais, é a diminuição das informações do cor do vídeo. Isso é possível separando as informações do vídeo em luminância e crominância. A crominância é menos sensível para a visão humana do que a luminância, e esta redução é pouco perceptível [20]. Para realizar essa redução, as informações do espaço de cor RGB do vídeo são transformadas para o espaço de cor YCbCr, e posteriormente as informações de Cb e Cr são reduzidas. A esta redução dá-se o nome de subamostragem de crominância.

Um vídeo com o espaço de cor YCbCr e amostragem 4:4:4 representa um vídeo com a mesma quantidade de informação de luminância (componente Y) e de crominância (componentes Cb e Cr), equivalente ao modelo RGB. Um vídeo que sofreu redução de crominância, sendo transformado no espaço de cor 4:2:2, por exemplo, representa um vídeo cujas informações de crominância foram reduzidas pela metade [20]. Na subamostragem 4:2:2 são necessários 16 bits para representação de cada pixel, enquanto no 4:4:4 são necessários 24 bits. Outra subamostragem utilizada no espaço de cor YCbCr, é a 4:2:0, onde para cada dois Y são associados alternadamente um de Cb e outro de Cr. Isto faz com que essa subamostragem necessite apenas de 12 bits por pixel, e quanto menor a quantidade de bits por pixel em cada subamostragem, maior a redução da crominância do vídeo. É

desejável que seja utilizada uma maior compressão, contudo isso pode acarretar em dificuldades na percepção anaglífica.

Afim de testar qual é o limiar dessa subamostragem, os vídeos estéreo anaglíficos deste trabalho foram subamostrados para 4:2:2 e 4:2:0, no qual os testes objetivos e subjetivos de qualidade foram realizados.

### 5. BASE DE TESTES

Apesar de a literatura reportar critérios de seleção de vídeos monoculares para compor uma base de testes [21] [22] para que desenvolvedores possam realizar testes de conformidade e de qualidade, não existem tais critérios para vídeos estereoscópicos. Em um trabalho anterior [18], havia sido proposto uma base de testes de vídeos estereoscópicos, contudo essa base possuía poucos vídeos e estes se encontravam em definição *standard*. Neste presente trabalho a base foi ampliada em seu número de amostras, assim como a definição dos vídeos (que aumentou para 960 x 540 pixels). Seguem abaixo os sete critérios para vídeos monoculares conforme os padrões do ITU-T:

1. Uma seqüência estressando cor;
2. Uma seqüência estática (com pouco movimento);
3. Uma seqüência estressando luminância;
4. Uma seqüência que contenha texto;
5. Devem existir várias seqüências;
6. Devem existir seqüências com movimentação;
7. Todas as seqüências originais não podem ter ruído.

De forma a acrescentar informações referentes aos vídeos estereoscópicos, foram avaliadas algumas informações de diversos planos (como as presentes nos critérios 1, 2, 3, 4 e 6) para que houvesse a variação determinada pelo critério 5. As informações analisadas estão descritas a seguir:

- **Enquadramento do Plano:** refere-se ao local e de que forma posicionar a câmera durante as gravações. Determinar o enquadramento significa pensar em qual área irá aparecer na cena e qual o ponto de vista mais indicado para que a ação seja registrada. Alguns tipos de enquadramento são:
  - **plano geral ou de conjunto:** mostra uma paisagem;
  - **plano médio:** mostra um ambiente;
  - **plano detalhe:** mostra uma parte do corpo humano ou objeto.
- **Movimento e Informações dentro do Plano:** com a estereoscopia, a movimentação torna-se presente em mais uma dimensão, o que traz muitas possibilidades tendo como base o vídeo bidimensional;
- **Cores e Luminância:** planos com maior e menor incidência de cores e com variação de luminância devem estar contemplados na base estereoscópica.

Para compor a base de teste, dois planos foram captados manualmente, com equipamento do Departamento de Artes e Comunicação da UFSCar. A produção destes planos se encontra detalhada em um relatório técnico [23]. Outros dois planos foram cedidos por um pesquisador<sup>2</sup>, no entanto o conjunto não possuía grande variação de informações, e para completar a base foi necessário procurar em diversos vídeos estereoscópicos<sup>3</sup>, alguns planos que tivessem variações das informações analisadas.

<sup>2</sup> <http://www.propp.ufms.br/gestor/titan.php?target=openFile&fileId=637>

<sup>3</sup> A grande maioria dos planos foi obtida em vídeos disponíveis no site <http://www.3dtv.at>

Para todos os planos dos pares estereoscópicos de vídeo, foi aplicado um filtro de correção de artefatos gerados pela compressão de dados. As codificações MPEG-2 e H264 (utilizadas pelas câmeras digitais) utilizam transformada DCT (*Discrete Cosine Transform*) e sua respectiva quantização, que acarreta a presença de artefatos de compressão ou *blockiness* [24]. Para minimizar esse ruídos, os vídeos foram processados no software VirtualDub<sup>4</sup> com o filtro *deblocking*, desenvolvido pelo MSU<sup>5</sup>. Este pré-processamento foi realizado para garantir que os planos estivessem sem ruídos, respeitando o critério 7. Os planos selecionados para a base estereoscópica estão relacionados na Tabela 2.

**Tabela 2. Planos dos vídeos estereoscópicos utilizados para testes de codificação**

Nº - Vídeo	Enquadramento e Movimento	Cores e Luminância	N. de Quadros
1 - ARV01	Plano Geral - Panorâmica com pouco movimento de transeuntes, em plano pouco profundo.	Verde e muitas variações; Alta luminância.	400
2 - CORR01	Plano Médio - Câmera e cenário estáticos; um transeunte se move do fundo para proximidades do primeiro plano.	Predominância de branco e azul; Média luminância.	260
3 - CRUZ01	Plano Geral - Movimentação de pessoas em primeiro plano e cidade no cenário de fundo, estático.	Presença de verde e algumas variações, com baixa luminância.	450
4 - DO01	Plano Geral - Movimento de câmera semi-circular, com objeto orbitando em primeiro plano.	Predominância de amarelo; alta luminância.	460
5 - DO02	Plano Detalhe - Câmera e cenário estáticos; movimentação de transeuntes no plano profundo	Presença de verde e algumas variações. Média luminância.	110
6 - DO03	Plano Médio - Movimentação na perspectiva diagonal; câmera estacionária.	Alta luminância, onde se faz predominante o azul claro celeste.	330
7 - DO04	Plano Detalhe - Câmera e cenário estáticos; movimentação de alta rotação na partida da hélice em primeiro plano	Alta luminância, onde se faz predominante o azul claro celeste e matizes de verde.	170
8 - DO05	Plano Geral - Objeto quase estático, oscilando no primeiro plano e cenário se movimentando lentamente ao fundo.	Presença de verde e algumas variações; alta luminância.	220
9 - DZ01	Plano Geral - Rotação do objeto em primeiro plano.	Média luminância, fundo em camafeu. Tons saturados entre verde e azul.	216
10 - DZ02	Plano Detalhe - Câmera em movimento dentro de cenário.	Predominância de vermelho/púrpura dessaturado, com baixa luminância.	130
11 - DZ03	Plano Detalhe - Rotação da câmera, pivotada no objeto em primeiro plano, relativo ao fundo.	Tons dessaturados, com predominância de amarelo e bege. Média luminância.	130
12 - DZ04	Plano Geral - Panorâmica no ambiente, seguido por movimentação da câmera.	Grande luminância e amplo espectro de cores entre amarelo, azul e vermelho	160
13 - FW01	Plano Geral - Câmera estática, com partículas se movendo em expansão centrífuga.	Baixa luminância. Predominância de amarelo dourado nas partículas	540
14 - FW02	Grande Plano Geral - Câmera estática, com partículas em explosão.	Baixa luminância. Presença de amarelo, verde azul, vermelho e tons mistos.	400

<sup>4</sup> O software Virtual Dub é um software livre que opera no sistema operacional Windows e realiza conversão de vídeo e áudio. <http://virtualdub.org/>

<sup>5</sup> O MSU Graphics&Media Lab Video Group é composto por um grupo de pesquisadores russos, da Universidade de Moscou. [http://compression.ru/video/deblocking/index\\_en.html](http://compression.ru/video/deblocking/index_en.html)

15 - HEI01	Grande Plano Geral - Câmera estática. Grafismo faz <i>zoom in</i> no primeiro plano	Média luminância, predominância do azul celeste, verde e vermelho-terra.	360
16 - HEI02	Plano Médio - <i>Zoom in</i> seguido de panorâmica.	Alta luminância, predominância do azul celeste, verde e tons médios	350
17 - HEI 03	Plano Geral - Panorâmica.	Alta luminância no segundo plano; predominância de vermelho e bege sépia	220
18 - HEI04	Plano Geral - Câmera estática; movimento de transeuntes na perspectiva.	Alta luminância no segundo plano e amplo contraste; predominância de tons azulados .	380
19 - MP01	Plano Médio - Câmera se movimenta e mantém os sujeitos do primeiro plano no quadro.	Média luminância; presença de azul saturado na iluminação	600
20 - OLD01	Plano Geral - Câmera estática, com objeto se movendo em perspectiva, da esquerda para a direita.	Alta luminância, com cores saturadas em detalhes pontuais (amarelo e vermelho).	470
21 - OLD02	Plano Geral - Câmera estática, com objeto se movendo em perspectiva, da direita para a esquerda.	IDEM	200
22 - OLD03	Plano Geral - Panorâmica.	Alta luminância com baixo contraste de cores. Presença de amarelo e vermelho	470
23 - OLD04	Plano Geral indo para Detalhe - Panorâmica.	IDEM	160
24 - RV01	Grande Plano Geral - Panorâmica com giro.	Alta luminância e tons de azul / ciano predominantes	130
25 - RV02	Grande Plano Geral variando em Plano Médio - Panorâmica seguida de movimentação da câmera.	Alta luminância e tons dessaturados de azul e verde. Detalhes em vermelho mais saturado	420
26 - RV03	Plano Geral - Panorâmica.	Média luminância, com predominância de tons azulados e detalhes pequenos em amarelo e vermelho saturados	340
27 - RV04	Plano Geral - Câmera com movimento enquadrando o plano de baixo.	Luminância média, com tons pastéis de verde predominantes.	160
28 - RV05	Plano Geral - Panorâmica suave , onde o objeto em primeiro plano se desloca relativo ao fundo.	Média luminância e cores dessaturadas.	190
29 - RV06	Plano Geral - Panorâmica com giro.	Alta luminância e tons dessaturados de verde, azul e marrom.	100
30 - SKY01	Plano Geral - Câmera com movimento relativo de translação ao objeto referencial, em segundo plano	Predominância de azul celeste; alta luminância.	200
31 - SKY02	Plano Geral - Câmera com movimento relativo de translação ao objeto referencial, em primeiro plano	Predominância de azul celeste e tons verdes; alta luminância.	500
32-TRAV01	Plano Geral - Câmera estática, com movimento de pessoas e veículos nos eixo de perspectiva da câmera (frente-fundo) e lateral (esquerda-direita).	Luminância média; cores saturadas em alguns objetos pontuais, ocorrendo o laranja, vermelho, amarelo, azul e verde.	600

Todos os 32 vídeos pré-processados possuem dimensões de 960x540 pixels e 24 quadros por segundo. A base de testes se encontra disponível na Web<sup>6</sup>.

## 5.1. Codificação

Para criar os vídeos estereoscópicos anaglíficos a partir do par de vídeos estéreo, foi desenvolvida uma aplicação em Java que realiza a leitura de dois arquivos AVI sem compressão (estes

<sup>6</sup> <http://200.136.217.194/videoestereo>

arquivos são pares estereoscópicos), e gera uma sequência de imagem BMP sem compressão. A ordem dos canais referente a combinação de cor foi explanada na seção 2, na qual o método ciano/vermelho utiliza as componentes  $R_2G_1B_1$  e o método magenta/verde utiliza as componentes  $R_1G_2B_1$ .

A partir das imagens geradas pela aplicação em Java, utilizou-se o software Virtual Dub para gerar os vídeos estereoscópicos e posteriormente para realizar a subamostragem 4:2:2 e 4:2:0. É importante ressaltar que nas operações realizadas de codificação de vídeo não foram utilizados CODECs de vídeo e os vídeos codificados não foram comprimidos.

## 6. MEDIDAS REALIZADAS

### 6.1. Medidas Subjetivas e Objetivas

Para medir o grau de interferência da subamostragem de crominância na percepção estereoscópica anaglífica foram realizados dois testes: um subjetivo, usando como medida o MOS (*Mean Opinion Score*) [17]; e outro objetivo, usando como medida o PSNR (*Peak Signal-to-Noise Ratio*) [17].

A medida subjetiva do MOS foi realizada para averiguar quanto a percepção de profundidade era afetada com a subamostragem dos vídeos anaglíficos. Seu cálculo foi realizado de acordo com as normas do ITU-T [21][22]. Inicialmente foram montadas estruturas de vídeo ABAB, onde A é o vídeo original e B é vídeo processado, e foi constatado que nos intervalos dos vídeos existem trechos de tela cinza, como mostra a Figura 3.



Figura 3. Estruturas de vídeo utilizadas para medida do MOS.

Essas estruturas de vídeo foram submetidas a um grupo de avaliadores composto de 32 pessoas, dentre os quais, indivíduos de ambos os sexos, entre alunos e professores da UFSCar com faixa etária de 18 a 50 anos, todos sem ter experiência alguma em avaliação de vídeos digitais e processamento de imagem e vídeo, conforme atestado por eles mesmos. O tempo de duração da avaliação foi de aproximadamente de 30 minutos e para esta situação foram utilizados óculos com filtros coloridos conforme citados na seção 2. Foi utilizado para a reprodução dos vídeos avaliados um monitor de 22 polegadas modelo LG W2252TQ, que possui taxa de brilho de 300 cd/m<sup>2</sup> e taxa de contraste de 10.000:1.

Tais avaliadores utilizaram uma tabela de pontuação cuja variação ocorre de 1 (pior qualidade) à 5 (melhor qualidade). Através do resultado do grupo de avaliadores, obteve-se o MOS, ou seja, a média de valores atribuídos aos vídeos pelos avaliadores.

Para tal procedimento (avaliação), foi necessário englobar os seguintes passos [22]:

(i) Teste de acuidade visual utilizando o teste de isihara<sup>7</sup>;

(ii) Treinamento e apresentação das possíveis imparidades encontradas;

(iii) Apresentação da sequência de vídeo.

Após a avaliação subjetiva, a próxima etapa constituiu-se de uma avaliação objetiva da qualidade dos vídeos. Como neste caso a qualidade dos vídeos está relacionada à cor, realizou-se o cálculo PSNR nas componentes de cor RGB, sendo que o PSNR avalia a quantidade de ruídos introduzidos na compressão tendo como base de comparação a imagem original, e seus valores mínimo e máximo de qualidade são 0 e 100 dB.

O PSNR foi medido em cada uma das componentes de cor dos vídeos codificados com a ferramenta MSU VQMT (*Video Quality Measurement Tool*<sup>8</sup>), ferramenta esta que calcula o valor do PSNR de cada quadro do vídeo codificado (comparando-o com o quadro correspondente do vídeo original), assim como uma média de PSNR de todos os valores.

Considerando os métodos de combinação de cor vermelho/ciano e verde/magenta, a média de PSNR das componentes azul e verde foi calculada para produzir os dados da componente ciano. Da mesma forma, a média do PSNR das componentes azul e vermelha foi calculada para produzir os dados da componente magenta.

A metodologia de análise do PSNR é distinta para cada método anaglífico, e envolve analisar suas médias (calculadas à partir do valor do PSNR de cada quadro do vídeo) para cada componente de cor. Valores muito distantes das médias de PSNR das duas componentes podem implicar em uma avaliação subjetiva na qual a profundidade não é percebida adequadamente [18].

### 6.2. Discussão dos Resultados

As Figuras 4 e 5 mostram o cálculo do MOS para as seqüências de teste ciano/vermelho enquanto as Figuras de 6 e 7 mostram o cálculo do MOS para as seqüências de testes magenta/verde. Nestes gráficos, o eixo horizontal representa o vídeo avaliado da Tabela 2, enquanto o eixo horizontal a média de MOS.

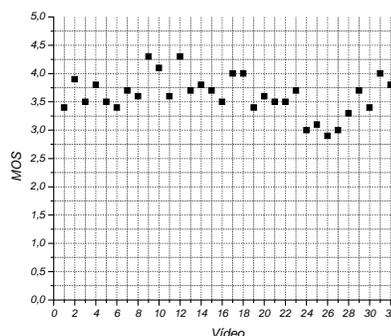


Figura 4. MOS dos vídeos vermelho/ciano codificados com a subamostragem 4:2:0

<sup>7</sup> [http://genetica.ufcspa.edu.br/seminarios%20monitores/2004/daltonismo\\_texto.pdf](http://genetica.ufcspa.edu.br/seminarios%20monitores/2004/daltonismo_texto.pdf)

<sup>8</sup> O software MSU VQMT foi desenvolvido pelo grupo de pesquisa MSU Graphics & Media Lab da Universidade de Moscou. [http://compression.ru/video/quality\\_measure/video\\_measurement\\_tool\\_en.html](http://compression.ru/video/quality_measure/video_measurement_tool_en.html)

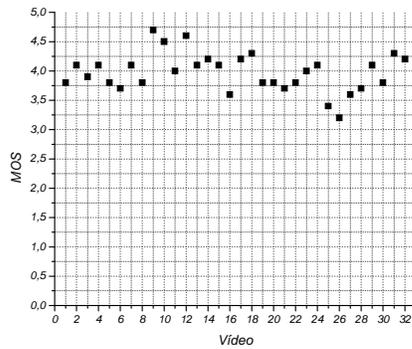


Figura 5. MOS dos vídeos vermelho/ciano codificados com a subamostragem 4:2:2.

Ao analisar o MOS dos vídeos codificados (Figuras 4 à 7), concluiu-se que os vídeos codificados com o espaço de cor 4:2:2, tem melhores resultados do que os vídeos codificados com o espaço de cor 4:2:0. Ao compararmos os vídeos vermelho/ciano com os verde/magenta, tendo como referência a análise das médias dos valores dos MOS de todos os vídeos, temos os dados apresentados na Tabela 3.

Os dados contidos na Tabela 3 apontam os vídeos magenta/verde codificado com a subamostragem 4:2:2 como os que possuem melhor percepção estereoscópica, dentre os vídeos analisados. Aliado a este fato, este método com a referida subamostragem, foi o único a ter as suas avaliações de MOS acima de 3,5 nas Figuras 4 à 7.

Na avaliação subjetiva, o índice de MOS 3 é definido como Regular, enquanto o índice 4 é definido como Bom [21]. Durante a avaliação subjetiva da percepção estereoscópica o índice 3,5 foi tido como um índice crítico para percepção. Valores abaixo desse índice indicaram vídeos cuja percepção estereoscópica não foi aceitável.

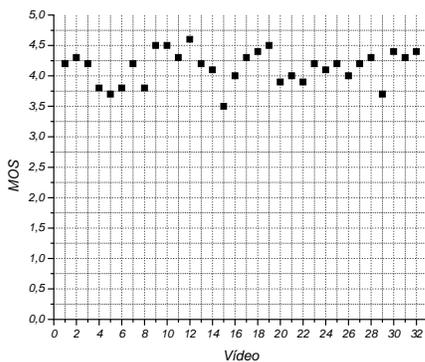


Figura 6. MOS dos vídeos verde/ magenta codificados com a subamostragem 4:2:0.

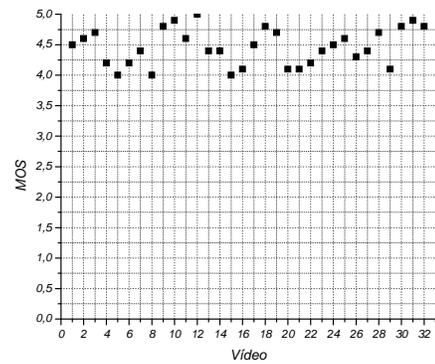


Figura 7. MOS dos vídeos verde/ magenta codificados com a subamostragem 4:2:2.

Tabela 3. Média dos valores de MOS dos vídeos codificados

Combinação	YV12 (4:2:0)	YV16 (4:2:2)
Vermelho/Ciano	3,61563	3,97187
Verde/Magenta	4,14063	4,45937

Os gráficos das Figuras 8 e 9 mostram os valores médios de PSNR para as componentes vermelho/ciano e as Figuras 10 e 11 para as componentes verde/magenta. Nestes gráficos, o eixo horizontal representa os 32 vídeos avaliados, enquanto que no eixo vertical se encontra o valor médio de PSNR.

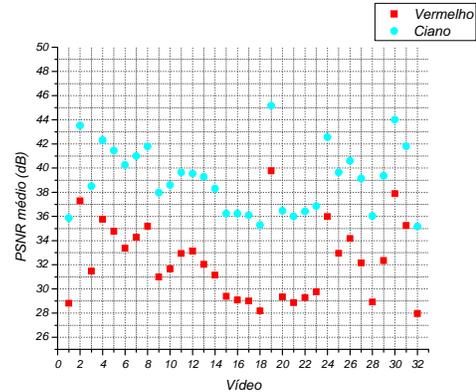


Figura 8. PSNR dos vídeos vermelho/ciano codificados com a subamostragem 4:2:0.

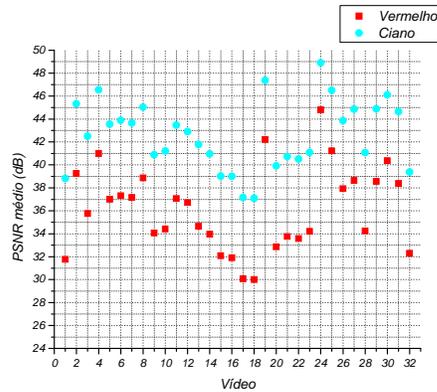


Figura 9. PSNR dos vídeos vermelho/ciano codificados com a subamostragem 4:2:2.

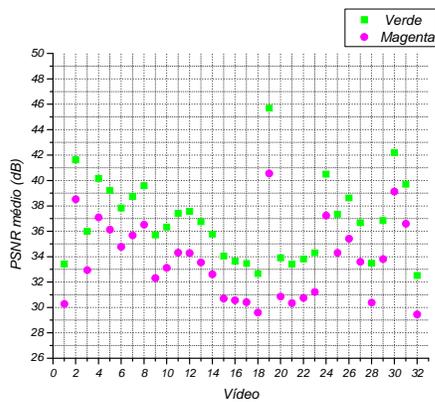


Figura 10. PSNR dos vídeos verde/ magenta codificados com a subamostragem 4:2:0.

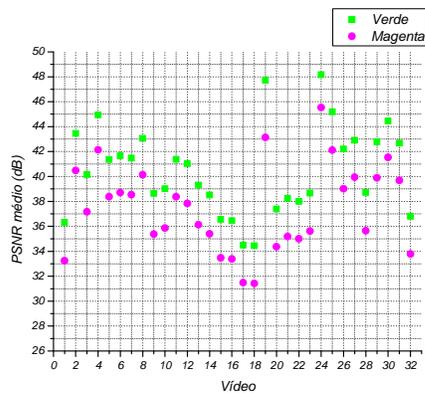


Figura 11. PSNR dos vídeos verde/ magenta codificados com a subamostragem 4:2:2.

Ao analisarmos as medidas objetivas, o PSNR dos vídeos codificados com o espaço de cor 4:2:2 obteve médias superiores ao PSNR dos vídeos com o espaço de cor 4:2:0. Dando continuidade a análise, ao verificarmos as medidas subjetivas, em especial os índices de MOS das Figuras de 4 e 5, vemos que as combinações de vermelho/ciano com subamostragem de cor

tiveram alguns vídeos com índices de MOS abaixo de 3,5, enquanto que para a combinação verde/magenta, observou-se nas Figuras 6 e 7 que apenas um vídeo da subamostragem apresentou índice de MOS igual a 3,5 (vídeo 15 da subamostragem 4:2:0).

Em relação ao cálculo do PSNR de cada método, é apresentada a Tabela 4, onde constatamos que o método vermelho/ciano obtém valores médios maiores de PSNR. Calculando as diferenças de PSNR das respectivas componentes de cor, obtem-se a Tabela 5, e desta informação observamos que quanto maior a diferença das médias de PSNR entre suas componentes de cor, como vemos nos dados apresentados na Tabela 5, pior a percepção subjetiva correspondente.

Tabela 4. Média dos PSNR dos vídeos codificados por subamostragem

	4:2:0	4:2:2
Ciano	39,12046	42,58279
Vermelho	32,28649	36,12640
Magenta	33,65208	37,43634
Verde	36,83681	40,50566

Tabela 5. Média das diferenças de PSNR dos vídeos codificados por subamostragem

	4:2:0	4:2:2
Vermelho/Ciano	6,83397	6,45639
Verde/Magenta	3,18473	3,07026

## 7. CONCLUSÕES

Ao verificarmos os dados da Tabela 3, constatamos que nas duas subamostragens aplicadas a diferença de MOS entre os métodos de combinação variou em torno de 0,5. Esse valor com relação ao método de avaliação subjetivo implica em 10% de acréscimo de qualidade, favorecendo o método magenta/verde em relação ao método vermelho/ciano.

Analisando as médias de PSNR da Tabela 5, verificamos que o método verde/magenta possui diferenças das médias de PSNR nas duas subamostragens da ordem de 3 dB, enquanto as diferenças das mesmas médias para o método ciano/vermelho é da ordem de 6 dB (mais do que o dobro do valor).

Ao cruzarmos os resultados das análises das Tabelas 3 e 5, verificamos que o método verde/magenta obtém resultados melhores que no vermelho/ciano para as duas subamostragens avaliadas. Ainda assim, se a comparação for realizada entre as subamostragens de crominância, a opção 4:2:2 apresenta MOS superior em 0,3 (cerca de 6%) em relação ao 4:2:0 nos dois métodos.

Considerando que o método 4:2:2 não teve nenhum valor de MOS abaixo de 3,5 e que apresentou os melhores resultados nas Tabelas 3 e 5, concluiu-se que o método de combinação anaglífica verde/magenta na subamostragem 4:2:2 se apresenta como a melhor alternativa, no escopo deste trabalho para a composição de vídeos estereoscópicos anaglíficos, tendo como objetivo a obtenção de uma melhor percepção da profundidade. Esse resultado pode ser utilizado para beneficiar diversos trabalhos relacionados que utilizam a técnica vermelho/ciano, substituindo-a pela verde/magenta para obtenção de ganho de qualidade da percepção estereoscópica.

Como contribuições deste trabalho têm-se uma base de testes para vídeos estereoscópicos e um software para a geração de vídeos estereoscópicos anaglíficos a partir de um par estéreo de vídeos.

Tendo em vista trabalhos futuros pretendem-se realizar testes e comparações com diferentes transformadas matemáticas, verificando como a etapa de quantização relacionada às transformadas matemáticas influencia na perda da percepção anaglífica, visando o estabelecimento de um método de codificação flexível que possa atender todas as demandas de aplicações de visualização estereoscópica atuais e futuras.

## AGRADECIMENTOS

Nossos agradecimentos à FAPESP pelo apoio à participação no evento, a Djalma Ribeiro Jr. e Pedro Dolosic, do DAC (UFSCar) que auxiliaram na montagem da câmera e nas captações estereoscópicas, e por fim a todos os cinegrafistas e ao animador que permitiram a utilização de seus planos na base de testes.

## REFERÊNCIAS

- [1] Adikari, A. B. B.; Fernando, W. A. C.; Arachchi, H. K.; Loo, K. K. *A H.264 Compliant Stereoscopic Video Codec*. Electrical and Computer Engineering, 2005. P. 1614-1617. ISSN: 0840-7789.
- [2] StereoGraphics Corporation. *Stereographics Developers Handbook: Background on Creating Imagens for CrystalEyes and SimulEyes*, 1997. Disponível em: <http://www.cs.unc.edu/Research/stc/FAQs/Stereo/stereo-handbook.pdf>. Acesso em: 10 abril 2010.
- [3] Zone, R. *Stereoscopic cinema and the origins of 3-D film, 1838–1952*. The University Press of Kentucky, 2007. ISBN: 0813124611.
- [4] J. Konrad and M. Halle. *3-D Displays and Signal Processing – An Answer to 3-D Ills?.* IEEE Signal Processing Magazine, Vol. 24, No. 6, Nov. 2007.
- [5] Perlin, K., Paxia, S. and Kollin, J. *An Autostereoscopic Display*. SIGGRAPH 2000 Conference Proceedings. New Orleans, Louisiana. July 23-28.
- [6] Perlin, K. et. al. *Recent Advances in the NYU Autostereoscopic Display*. Proceedings of the SPIE, v. 4297, São José, California, 2001.
- [7] Mendiburu, B. *3D Movie Making Stereoscopic Digital Cinema from Script to Screen*. ISBN: 978-0-240-81137-6. Ed. Focal Press, Elsevier, 2009.
- [8] Suppia, A. O. *Monstro Brasileiro Revive em 3D*. Revista Cultura e Ciência, ano 59, n.2, p.57-59, 2007. ISBN 85-0009-6725.
- [9] Sony Group. *Sony's 3D TV*. Disponível em : <<http://www.sony.net/united/3D/#technology/3dtv/>>. Acesso em: 01 abril 2010.
- [10] Smolic, A.; Mueller, K.; Merkle, P.; Kauff, P.; Wiegand, T. *An Overview of Available and Emerging 3D Video Formats and Depth Enhanced Stereo as Efficient Generic-Solution*.

Proceedings of the 27th conference on Picture Coding Symposium. Chicago, 2009.

- [11] Lipton, L. *Stereo-Vision Formats for Video and Computer Graphics White Paper*. Disponível em <[http://www.cours.polymtl.ca/inf6802/stereo/body\\_stereo\\_format\\_s.html](http://www.cours.polymtl.ca/inf6802/stereo/body_stereo_format_s.html)>. Acesso em: 10 abril 2010.
- [12] Matsuura, F.; Fujisawa, N. *Anaglyph Stereo Visualization by the Use of a Single Image and Depth Information*. Journal of Visualization archive, v. 11, n. 1, p. 79-86, janeiro de 2008.
- [13] Menezes, M. S. et al. *Estereoscopia Aplicada à Neuroanatomia - Estudo comparativo entre as técnicas de filtro de cores e de polarização*. Arquivo de Neuro-Psiquiatria, v. 60, n. 3B, São Paulo, 2002.
- [14] Martin, S., Rubio, R. *Parallax cues in the design of graphics used in technical education to illustrate spatial problems*. Computer & Education. Elsevier Science, 2009. Vol 52, n 2, pp. 493-503.
- [15] Hartman, N. W., Bertoline, N. W. *Spatial Abilities and Virtual Technologies: Examining the Computer Graphics Learning Environment*, iv, pp.992-997, Ninth International Conference on Information Visualisation (IV'05), 2005.
- [16] Andrade, L.; Dolosic, P.; Ribeiro, D. J. *Vídeo digital estereoscópico: aspectos tecnológicos e implicações na narrativa*. In: Congresso SET - Broadcast&Cable, 2009, São Paulo-SP. Revista de Radiodifusão. São Paulo - SP : EMBRASEC - EDITORA & EVENTOS, 2009. v. 03. p. 338-345.
- [17] Gonzalez, R. C., Woods, R. E. *Digital Image Processing. Third Edition*, Ed. Prentice Hall, 2008. ISBN: 013168728X.
- [18] Andrade, L., Goularte, R. *Percepção Estereoscópica Anaglífica em Vídeos Digitais Comprimidos com Perda, Webmedia – Brazilian Symposium on Multimedia and the Web*, Fortaleza, v. 1, n.1 , pp. 226-233. 2009.
- [19] Dubois, E. *A projection method to generate anaglyph stereo images*. Proc. IEEE Int. Conf. Acoustics Speech Signal Processing, vol. 3, Maio de 2001, p. 1661-1664.
- [20] Poynton, C. *Digital Video and HDTV Algorithms and Interfaces*. Morgan Kaufmann Publishers, 2003.
- [21] ITU-T *Recommendation J.247: Objective perceptual multimedia video quality measurement in the presence of a full reference*. International Telecommunication Union, 2008.
- [22] ITU-R, *ITU-R Recommendations BT-500.11 methodology for the subjective assessment for the Television Pictures*. International Telecommunication Union, 2002.
- [23] Andrade, L.; Dolosic, P.; Goularte, R. *Construção de uma Base de Vídeos Estereoscópicos*. São Carlos: Relatório Técnico, Nr. 351, ICMC-USP, 2010.
- [24] Ebrahimi F.; Chamik M.; Winkler S. *JPEG vs. JPEG 2000: An Objective Comparison of Image Encoding Quality*. Proceedings of SPIE, 2004.