Reversing Anaglyph Videos Into Stereo Pairs

Matheus Ricardo U. Zingarelli
Universidade de São Paulo
Instituto de Ciências Matemáticas e de Computação
Av. Trabalhador São Carlense, 400
São Carlos, São Paulo, Brasil
zinga@icmc.usp.br

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

There is a diversity of strategies for stereoscopic video coding, commonly known as 3D videos. Each strategy focuses on one type of 3D visualization format, which may lead to incompatibility or depth perception problems if one strategy is applied to another format. A generic approach may be based on the proper coding of the stereo pair – every visualization format knows how to deal with stereo pairs – however, to keep the stereo pair, even a coded one, demands a high data volume. This paper proposes a method for reversing an anaglyph video back into a stereo pair. This way, a coder may convert a stereo pair into an anaglyph video, reducing the data volume by half, at least. Moreover, a correspondent decoder may use the proposed reversing method to restore the stereo pair, ensuring the visualization independence of the coding method. Tests showed that both conversion and posterior reversion resulted in videos with good quality obtaining an average of 63.09% of compression and 34,52dB of PSNR.

Categories and Subject Descriptors

E.4 [**Coding and Information Theory**]: Data compaction and compression; I.4.2 [**Image Processing and Computer Vision**]: Compression (Coding) – *approximate methods*

General Terms

Algorithms, Performance, Standardization.

Keywords

Anaglyph video, stereo video coding, stereoscopy, digital video coding.

1. INTRODUCTION

Stereoscopic videos, commonly known as 3D videos, are formed by a pair of videos – called stereo pair (right eye, left eye) – and are reproduced in a way that gives a depth perception for a person watching them, mimicking the human stereo vision [4] (Section 2). Over the last few years, there's been an increased boost of 3D content production by the movie industry, largely due to the acceptance and expression of public interest for this technology. Besides that, 3D technology is being gradually incorporated at homes in forms of 3D television [15], cell phones [8] and video games [12] with each device supporting different kinds of visualization. Consequently, new techniques for capturing, coding and playback modes of stereoscopic videos are emerging or being improved in order to optimize and integrate this technology with the available infrastructure.

In the stereo video content production field, new cameras were developed for recording two views of the same scene, in order to produce the stereo pair, with the possibility of also generating a depth map of the scene that can be used to create new views [7]. There are also techniques developed for converting 2D content into 3D [18]. In the field of visualization, we have techniques for viewing stereo content with the support of specific glasses – anaglyph stereoscopy [9], polarized light [9] and shutter glasses [16] –, and also autostereoscopic displays, or glasses free, that allow us to view 3D content without wearing any type of glasses or specific devices [14].

In spite of the advances made in the field of visualization and representation of stereoscopic videos, it's noticeable that advances in the coding field are slower. On one hand, we have Lipton's Method [10], which describes formats for stereoscopic video representations, being the stereo pair stored in a single video container, without compression, with double of data than a regular 2D video stream. Since Lipton's Method keeps the stereo pair, it can be used by any visualization system – it is generic. On the other hand, we have what we call "adapted methods" that use well-known compression techniques, like MPEG-2 or H.264, to reduce the amount of data to be transmitted. However, such techniques are only adapted to work on stereoscopic videos, lacking of a standard compression technique specific designed for this kind of video. Moreover, each adapted method is designed for a particular visualization system, which brings two problems: they may not be compatible for all formats and types of stereo visualization [14], and since lossy compression is used, depth perception may be compromised in some cases, especially with anaglyph videos [2][3]. With that said, one can realize that there's a lack of a generic method specific for coding stereoscopic videos, compatible with different types of 3D visualization and providing good quality without loss of depth perception.

Related work has been done looking for compression of stereo video pairs without significant quality loss. Results like [3] demonstrate that it's possible to develop coding methods to reduce the data volume having no depth perception loss and, more important, being independent of the visualization method. In spite of the reduction of data volume achieved by those coding methods, the compression rate still remains low since they keep the stereo pair. A straightforward solution is the conversion of the stereo pair into a single anaglyph video stream (Section 4.1), resulting in a stereo video with higher compression rate, superior quality and, obviously, compatible only with the anaglyphic visualization method.

This way, in order to keep the generality of such coding methods, making them also compatible with other kinds of visualization, anaglyph to stereo pair reversion methods are needed. A reversion method is not trivial though, since the anaglyph conversion causes loss of color components in the stereo pair, and they must be retrieved somehow. This paper demonstrates that this reversion is possible and proposes a technique based on a compressed chroma sub-sampled color index table. With this technique we were able to achieve 63.09% of compression and to create a reversed stereo pair with no loss in depth perception.

This paper is organized in the following sections. Section 2 presents details about stereo vision and depth perception needed for a better understanding of the proposed technique. Section 3 describes related works about stereo video coding and formats. Section 4 presents the technique we propose to tackle the problem of reversing an anaglyph video. Section 5 presents the experimental tests and results of using our technique. Finally, in Section 6 we present our conclusions and future work.

2. STEREO VISION AND DEPTH PERCEPTION

Our eyes are approximately 6.5 cm distant from each other, move together in the same direction and each one has a limited viewing angle. By presenting themselves in different positions, each eye sees a slightly different image [4]. For these reasons it was expected that when we look at an object, we would saw two images and not just one. However, the brain takes charge of calculating information from relative distances from objects and to interpret these two images, resulting in production of a single image, phenomenon known as stereoscopy. The main stereoscopic information are stereopsis, (binocular) disparity and parallax [16].

The stereopsis is responsible for the depth sensation that we have between objects and is obtained due to binocular disparity. Thus, the mandatory requirement to obtain stereopsis is to use both eyes. With this information we feel objects closer or farther away. It is explored in 3D movies to give us the impression that objects are "bouncing off the screen" [16].

The binocular disparity is the difference in distance between the positions of the image formed on each retina. This is best understood through the following example: observe an object in front of you and place your thumb between your eyes and the object. When we focus on the thumb, i.e., it is at the point of convergence of the two retinas, the object is past the point of convergence (farther), appearing as doubled (Figure 1 (A)). This happens because the images off the focal point are being formed at different locations in each retina. The disparity is the distance between these two duplicate images. The same happens if we put our focus on the object (Figure 1 (B)).

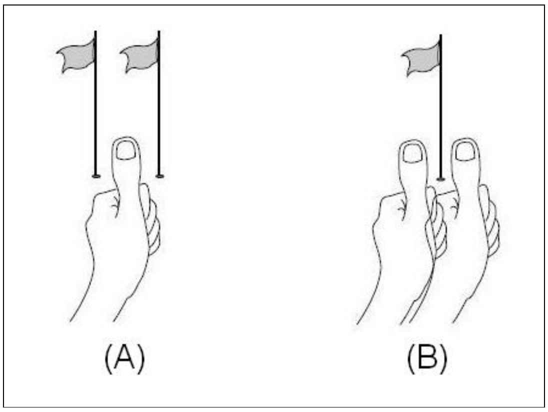

Figure 1 – Retinal Disparity [16].

Directly related to the disparity concept (obtained in the retinal image) we have the parallax [16][11], which is the distance between corresponding points in the images projected on a monitor. With the values of parallax, it is possible to give a different point of view of the same image to each eye, resulting in the formation of the disparity, and it therefore has the stereopsis effect. An easy way to calculate the parallax between two points is superimposing an image to another and measuring the distance between the same points in each image. It is because of parallax that, for example, when watching an anaglyph video without the proper glasses, we see parts of the image as duplicated and overlapped.

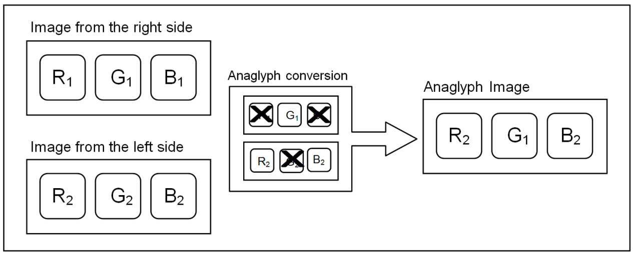

Figure 2 - Anaglyph Conversion to create a green-magenta anaglyph.

This way, stereoscopy is based on methods that present to an observer a pair of flat (2D) images from the same object, each one providing each eye with a slightly different perspective – the stereo pair. The stereo pair causes retinal disparity, then stereopsis, which, in turn, provides depth perception. As previously mentioned at Introduction, there are four main stereoscopic visualization methods: anaglyphic, polarized light, shutter glasses and autostereoscopic. The focus of this work is the anaglyphic method, which will be briefly described in the following.

The anaglyphic method is an important low cost and simple way of coding a stereoscopic pair to be visualized using proper glasses. In this method, we simply remove some color components from each image and then merge them into a single one. Afterwards, by using glasses with lenses that mimics the color components removed, thus acting as filters, images are separated again and each eye sees only one of them, leading to the binocular disparity and resulting in the stereoscopic effect.

Figure 2 illustrates the anaglyph conversion. The RGB channels of each image are combined in a way to have information from both images (the left eye and the right eye) [11]. Let the stereo pair be formed by $R_1G_1B_1$ (the right eye image) and $R_2G_2B_2$ (the left eye image). The conversion takes advantage of the green channel (G_1) of the right eye image and channels R_2 and B_2 of the left eye image, forming a third image (the anaglyph image) with information of both images of the stereo pair: $R_2G_1B_2$. As the combination of red (R_2) and blue (B_2) results in the magenta color, the anaglyphic method illustrated in Figure 2 is known as the green-magenta. Other possible methods are: red-cyan ($R_1G_2B_2$) and blue-yellow ($R_2G_2B_1$).

Dubois [6] and Andrade & Goularte [3] reported that green-magenta method presents better results for anaglyphic visualization than the other methods. Therefore, in this work, we use the green-magenta color combination.

Besides being simple, the anaglyphic method also doesn't require expensive or complex equipment to be executed or visualized, and the compression rate obtained is high, since only one video stream is transmitted/stored. The other visualization methods may have to keep the stereo pair, i.e., two video streams. Such advantages show that the anaglyphic method can be a potential candidate to be used in a generic stereoscopic coding process. Instead of sending the stereo pair, we could transform it into an anaglyph, reducing the data to be transmitted by half. This, however, comes with a cost: we are losing color components, and as mentioned in Section 1, that leads to a problem of recreating the stereo pair using a reversion method.

3. RELATED WORK

The MPEG group has well-known standard methods for video codification, even with extensions for stereoscopic videos, like MPEG-2 Multiview Profile [14]. However, there is no standard codification method specific for stereoscopic videos only. With that, different authors have created different coding strategies, each designed to attend requisites of one or another type of 3D system, which makes the implementation device-dependent and may result in incompatibility of content between different systems.

Among the different strategies, we can cite Lipton's Method [10] which describes several ways for presentation of a stereo pair of videos with concern of having little or no modifications in the hardware already available. They can be classified into two groups: field sequential scheme, in which right and left fields are alternated in sequence and the user, with proper glasses that synchronizes with the display, views only one of them on each eye. And pixel sequential scheme, with the above-and-below and side-by-side formats, in which the right and left subfields are united in a single field either horizontally or vertically.

Smolic et al. also stated in [14] the diversity of stereoscopic video formats, each directed to a specific system, thus requiring different implementations and structures. The authors classify these formats based on the number of video signals (called views), order of complexity and types of data involved, resulting in six classes. Conventional Stereo Video (CSV) is the simplest one, similar to Lipton's Method. Multiview Video Coding (MVC) is an extension to when more than two views are used. Video plus depth (V+D) is a format with more complexity, in which a depth map is sent together with the video signal to create the stereo pair, also enabling the possibility to generate a limited number of other views. These first three classes can be implemented using available video codecs. For advanced video applications like autostereoscopic televisions, the next three classes are used.

Multiview plus depth (MVD) is a combination of MVC and V+D properties, which means that multiple views and multiple depth maps are sent. The next format is layered depth video (LDV), in which besides a video signal and depth map, it is also sent a set of layers and associated depth maps used to generate virtual views. Finally, the last format is called depth enhanced stereo (DES), proposed by the authors as a generic 3D video format. It's an extension of the CSV, with the additional of depth maps and layers, providing compatibility among different formats, since each one uses only the types of data needed. Even though DES is designed to be a generic format, there are two major drawbacks that need a deep study: depending on which format data will be represented, it may be necessary the storage of a great amount of data to hold both video signal and additional depth maps and layers. We also have an increase in the system's complexity and errors that may arise from depth calculations.

Notice that in Lipton's Method both videos from the stereo pair are stored, resulting in a video file twice as big as a normal 2D video file, while Smolic et al. describe formats in which the stereo pair are not necessarily needed, what we call "adapted methods". Some authors study compression techniques to reduce the amount of data to be transmitted in these adapted methods, either using well-known compression techniques or designing new ones [17]. Vetro [19] performed a survey over the different formats and representation of stereoscopic and multiview videos, with their corresponding compression techniques and several types of displays in which they can be visualized. This survey clearly demonstrates the challenge in creating a generic method for stereoscopic video representation and coding: each surveyed method has specific types of compression, representation, storage, and plays only specific types of media and displays (visualization methods).

Andrade & Goularte [2] [3] studied how the usage of well-known lossy compression techniques on stereoscopic videos might affect quality and depth perception in different types of stereo visualization. They showed that the data lost during compression compromises depth perception in the anaglyph stereoscopy, and discovered suitable parameters for color space reduction, used together with Wavelet transform and quantization, that could compress stereoscopic videos with good quality and no depth perception loss regarding the anaglyph stereoscopy. In that work authors demonstrate the viability of a generic stereoscopic coding method, however, their method stores a stereo pair of videos, lowering the compression rate.

Thus, through related work, one can conclude that there is a lack of stereoscopic coding methods that: a) are independent of visualization methods, making the coding process generic enough to be possible to achieve stereo depth perception by the means of any visualization method; b) achieve high compression rates preserving image quality and depth perception, while keeping the visualization independence.

4. THE ANAGLYPH REVERSION TECHNIQUE

The reversion process is not trivial because during the anaglyph video generation some information is discarded. The stereo pair has six color channels (Figure 3): three (R_1, G_1 and B_1) from the right eye video/image, and three (R_2, G_2 and B_2) from the left eye video/image. When generating the anaglyph video, three channels are discarded, one channel from one video and two channels from the other video (at any combination). The challenge here is to recover lost information without significant compromise of depth perception and achieving good compression rates.

In order to recover the stereo pair from an anaglyph video, a first attempt is to store the discarded color information into some data structure, let's call it "Color Index Table". Following Figure 2 example, this table will be formed by color information from channels R_1, G_2 and B_1 and stored together with the anaglyph video $R_2G_1B_2$. This way, a decoder will have all the needed information to rebuild the stereo pair. In spite of this approach to keep color quality (it will preserve the color data), it does not present any compression advantage.

As only the color information is required in order to build the Color Index Table, a better approach is to use a color space conversion, from RGB to YC_bC_r [13]. This way, it is possible to separate luminance information (Y) from color information (C_b and C_r), using just C_b and C_r to compose the table. Moreover, C_b and C_r channels may be sub-sampled, reducing even more the amount of data needed to be stored in the table.

There are some possible color (sub) sampling combinations, presenting different tradeoffs between compression and color fidelity [13]. The 4:4:4 method is the best in quality, but the worst in terms of information reduction. The 4:1:1 method is exactly the opposite. Since colors greatly influence the anaglyphic method [2], Andrade & Goularte [3] have developed a study concluding that the 4:2:2 chrominance sub-sampling method offers a good tradeoff without affecting depth perception. Therefore, in this work, we use the 4:2:2 method.

The next three sections (4.1, 4.2 and 4.3) present how the anaglyph video is produced using chrominance sub-sampling in order to build the Color Index Table, how this table is used in order to reverse the anaglyph video back into a stereo pair, and a discussion about the method used.

4.1 The color index table

The Color Index Table is built following 4 steps, depicted in Figure 3: (I) creation of a green-magenta anaglyph video from the uncompressed stereo pair; (II) creation of another anaglyph with the remaining color components, which we call "complementary anaglyph"; (III) conversion of the complementary anaglyph from RGB to YC_bC_r color space using 4:2:2 sub-sampling and (IV) compression and storage of C_b and C_r components to create the table.

The production of a green-magenta anaglyph video from a stereo pair follows the scheme depicted in Figure 2. Each image from the stereo pair, named right image and left image, has its R, G and B color components separated. Then a new image is created, whose green color component is from the right image and red and blue (magenta) color components are from the left image ($R_2G_1B_2$). Doing this for every frame will result in the anaglyph video. Here, for the sake of clarity, we will explain the entire process using images as examples.

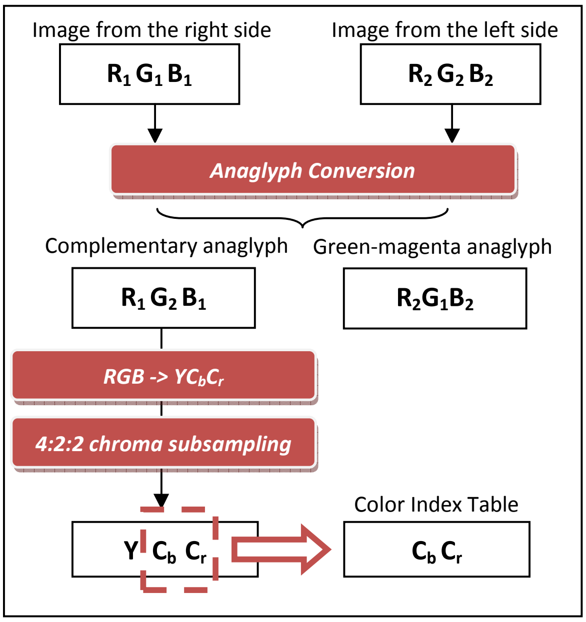

Figure 3 – Anaglyph conversion of a stereo pair of images and creation process of the Color Index Table

Figure 3 illustrates the creation of the green-magenta anaglyph image, which is kept apart meanwhile, and a complementary anaglyph image. This last one is generated by the color components not taken at first, that is, the red and blue color components from the right image and the green one from the left image ($R_1G_2B_1$).

Then, the complementary anaglyph is converted from RGB to YC_bC_r color space, using Equation 1 [5] to calculate color conversion. Equation 1 is an ITU-T recommendation.

```
Y  ← 0.299 * R + 0.587 * G + 0.114 * B
Cb ← (B - Y) * 0.564 + delta
Cr ← (R - Y) * 0.713 + delta,

                {128 for 8-bit images,
where delta =   {32768 for 16-bit images,
                {0.5 for floating-point images
```

Equation 1 – RGB to YC_bC_r conversion

The result is a 4:4:4 YC_bC_r image, and the complementary anaglyph is discarded. The 4:4:4 YC_bC_r image has its Y component discarded and its C_b and C_r components are 4:2:2 sub-sampled. It means that, for each 12 pixel samples (4:4:4) of the original image, 8 were discarded: 4 from the Y component, 2 from the C_b component and 2 from C_r component. This two sub-sampled components form the Color Index Table (Figure 3), which will be stored together with the anaglyph video. It should be noticed that, in spite of the Color Index Table being an overhead, it has only 33% of the complementary anaglyph data volume and, more compression will be achieved after applying Huffman lossless compression technique - Section 5 presents the results.

4.2 Anaglyph reversion

The anaglyph reversion consists of recreating the stereo pair using the Color Index Table and the anaglyph image, as depicted in figure 3. In order to obtain the stereo pair, we need the color components lost during the anaglyph conversion. These can be extracted from the Color Index Table by recovering the luminance component (Y) and applying a formula to convert from YC_bC_r to RGB. The luminance component can be calculated from the green-magenta anaglyph image using Equation 1. The conversion from YC_bC_r to RGB color space can be done by using Equation 2 [5]. Again, this formula is an ITU-T recommendation. Since we sub-sampled the image during the anaglyph conversion, we need to first duplicate each chrominance component to every 4 samples of luminance component and then apply Equation 2.

```
R <- Y + 1.403 *(Cr - delta)
G <- Y - 0.344 *(Cr - delta) - 0.714 *(Cb - delta)
B <- Y + 1.773 *(Cb - delta),

                {128 for 8-bit images
where delta =   {32768 for 16-bit images
                {0.5 for floating-point images
```

Equation 2 – YC_bC_r to RGB conversion

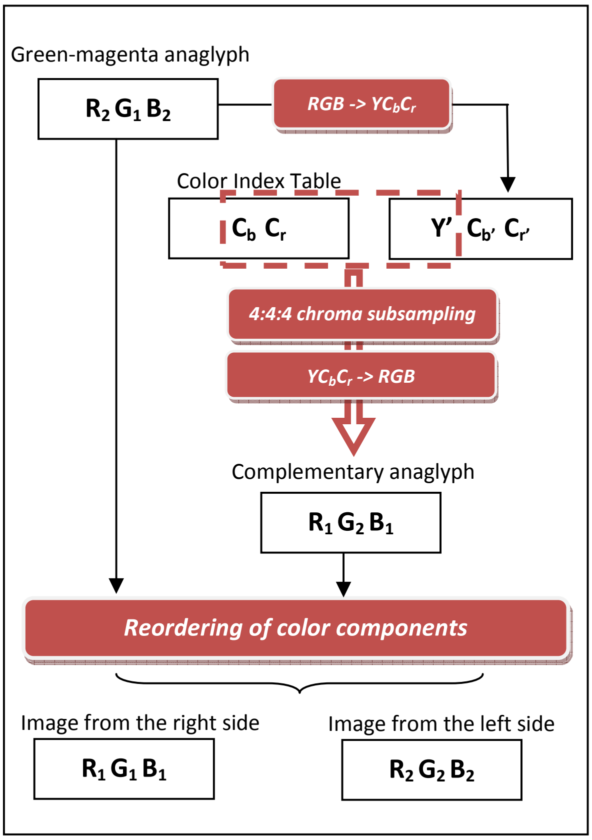

Figure 4 - Anaglyph reversion using a Color Index Table

Now that we have the missing color components, we can recreate the complementary anaglyph and then just need to reorganize the color components pertinent to each image of the stereo pair. It means that we take the red and blue color components from the green-magenta anaglyph and the green from the complementary anaglyph to recreate the left-eye image of the stereo pair. Likewise, with the green color component from the green-magenta and the red and blue from the complementary, we are able to recreate the right-eye image. Figure 4 summarizes the reversion process.

4.3 Method discussion

An important question, worth to discuss here, is the reason behind the use of the complementary anaglyph in order to build the Color Index Table.

The goal of the index is to store color information lost during the anaglyphic conversion. Since color information may take space, our approach is based on using chrominance sub-sampling in order to reduce the amount of data needed to be stored.

A first method one may think, trying to keep color quality at the best possible, is to apply the YC_bC_r conversion and sub-sampling directly to the stereo pair storing the chrominance components. This method will store 4 chrominance components together with the anaglyph image/video, instead of the two in our approach. In addition, in the reversion phase (Section 4.2), the Y component will have to be restored based on the anaglyphic version of the stereo pair (the stereo pair does not exist anymore), which has different information from those originally used to generate Y. This will result in distortions in the recovered image/video.

A second attempt is not to use a complementary anaglyph and to apply the YC_bC_r conversion and sub-sampling to the anaglyph image/video, generating the Color Index Table as described in Section 4.1. At the reversion phase the Y component may be properly restored based on the anaglyph image/video (Section 4.2) since it was also generated based on the anaglyph image/video. However, the C_b and C_r color components retrieved in this way have color information coded mostly from just one of the images in the stereo pair. For example, C_b and C_r may represent colors from a $R_2G_1B_2$ anaglyph image, which means there is no color information about R_1, G_2 and B_1 components in order to properly recreate the stereo pair. This will also result in distortions in the recovered image/video.

By using the proposed approach (Figure 3), the anaglyph image has color information that came from three of the six stereo pair color channels: R_2, G_1 and B_2, and the Color Index Table (C_b and C_r sub-sampled color components) carries color information obtained from the complementary anaglyph ($R_1G_2B_1$ in Figure 3). So, as we have color information that came from all the six RGB color channels of the stereo pair, we have conditions to properly recreate the stereo pair. This is done, as explained at Section 4.2, applying RGB-to-YC_bC_r conversions in order to recreate the complementary anaglyph image.

This way, we have R_1, G_2 and B_1 color channels from the complementary anaglyph and R_2, G_1 and B_2 color channels from the anaglyph image. Reordering the color channels, we get back the six channels of the stereo pair without depth distortions in the recovered image/video.

5. EXPERIMENT AND RESULTS

In our experiment, we focused on evaluate quantitatively the quality of stereo images outputted by our anaglyph reverse technique, based on 3 criteria: brightness, saturation and contrast, summarized in Table 1. Brightness has to do with intensity levels of luminance. Since each image from the stereo pair has a different point of view of a same scene, depending on the environment that they were captured, different intensity of brightness may appear. Saturation means the purity of a color, i.e. how much of white light there is in this color, where low saturation means higher amount of white light and vice-versa. Last criterion is contrast, the difference between adjacent colors in the image. The more two colors are different, the higher is the contrast, which allows better visualization in the details of an image. We've used a set of 32 stereo images classified between these criteria. Since the criteria are not mutually exclusive, one of more of them may appear in the same image. Images were extracted from a test database of stereoscopic videos created in [1] [3]. The database is available online and can be visited in http://200.136.217.194/videostereo/.

Our analysis on the quality of the stereo images reverted from our technique were based on calculating the PSNR – Peak Signal-to-Noise Ratio – of each pair of images: the original and the reverted one. The PSNR is a metric widely used to evaluate how similar is an image compared to another one [20]. It makes a pixel-by-pixel comparison and returns a value measured in decibels (dB), in the range of 0 to 100, with 0 meaning no similarities and 100 meaning total similarity.

Each PSNR was calculated using a free verion of a software called MSU VQMT[1] (Video Quality Measurement Tool). This software implements several evaluation metrics for image and video assessment, PSNR being one of them. For each criteria, we've calculated the PSNR of each color component of a pair of image in the RGB color space and then the average of the three results.

Table 1. Criteria used in the evaluation of the anaglyph reversion technique

Criterion	Types of images
1. Brightness	Images with high levels (brighter) or low levels (darker) of luminance
2. Saturation	Images with the presence of one predominant color and different levels of purity.
3. Contrast	Images with great or little variety of colors.

5.1 Results

For brightness, we've tested 18 images with high, medium and low levels of luminance. The average of PSNR among all images was 33,70dB, with a maximum of 37,39 dB belonging to an image of high luminance contrast – very brighter in some regions and darker in others –, and a minimum of 29,38 dB belonging to an image with high levels of luminance and little contrast. Difference in the values of PSNR between R, G and B color components on each image stayed on the range of 2,55 dB. The average of the compression rate was of 63.12%. Figure 5 shows the PSNR values obtained for each image.

For saturation, we've tested 20 images with different levels of color purity for the predominant colors. The average of PSNR among all images was 38,43 dB, with a maximum of 39,62 dB belonging to an image of medium contrast and presence of

[1] Available at http://compression.ru/video/quality_measure/index_en.html

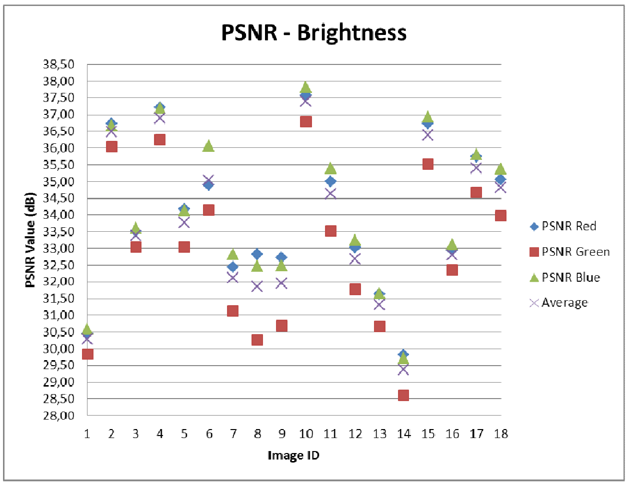

Figure 5 - PSNR values of images from the brightness criterion.

saturation levels of green and brown colors, and a minimum of 30,28 dB belonging to an image with the predominance of color green in different levels of saturation. Difference in the values of PSNR between R, G and B color components on each image stayed on the range of 2,55 dB. The average of the compression rate was of 63.61%. Figure 6 shows the PSNR values obtained for each image.

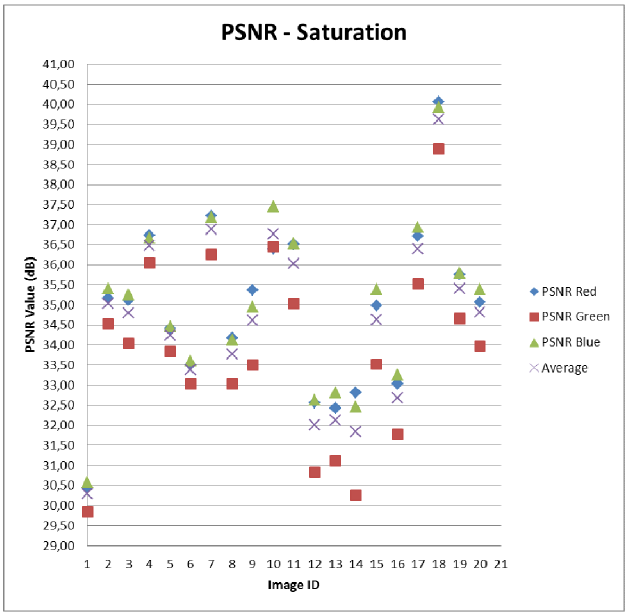

Figure 6 - PSNR values of images from the Saturation criterion.

For contrast, we've tested 17 images with great, medium and little variety of colors. The average of PSNR among all images was 33,05 dB, with a maximum of 39,62 dB belonging to the same image tested for saturation, and a minimum of 29,38 dB belonging to the same image tested for brightness. Difference in the values of PSNR between R, G and B color components on each image stayed on the range of 2,05 dB. The average of the compression rate was of 64.17%. Figure 7 shows the PSNR values obtained for each image.

From the results, we can observe that the PSNR value and the compression rate were similar in the three criteria defined, with a general average of 34,52 dB of PSNR and 63.09% of compression. We can also observe that the difference between the PSNR value of each RGB component of an image did not exceed 2,55 dB, which, accordingly to [2] is an acceptable value, in which the depth perception is not affected – difference of values greater than 5 dB are prohibitive, since it affects depth perception.

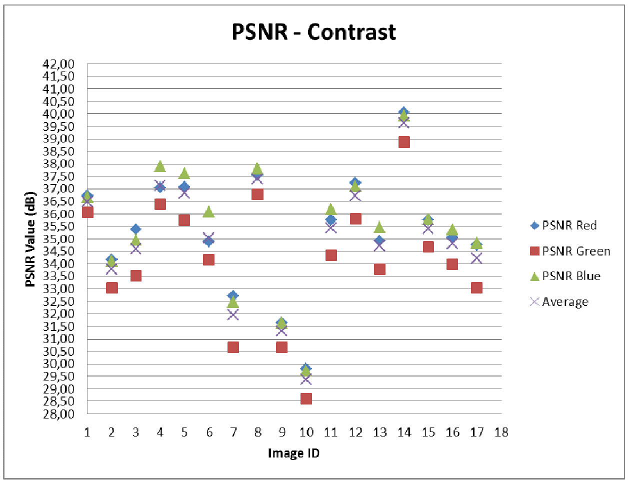

Figure 7 - PSNR values of images from the contrast criterion.

5.2 Image quality and depth perception

In our experiment, we took an anaglyph image and used our reversion technique to obtain a stereo image. This stereo image was compared to the original image and we calculated its PSNR. Even though the average of PSNR obtained was low, when comparing qualitatively both the reverted and original image, one can notice only a few differences between then, mainly caused by the pixel displacement that is present between the two images of the stereo pair. That means that even with a low PSNR value, the quality of the reverted image is still good. The PSNR value was affected by the conversion between RGB and YC_bC_r color spaces, which results in float values. Our technique was implemented in a way that each pixel is stored in an unsigned char variable, thus we have rounding errors involved in the process. As a future work, we will be reviewing our implementation, in order to achieve better PSNR results.

We took another final experiment regarding the reverted images obtained by our technique. Each one was converted into a green-magenta anaglyph and we compared qualitatively this anaglyph with the anaglyph formed by the original image. The anaglyph from the reverted image presented good quality and depth perception was not lost, which proves that the maximum difference of 2,55 dB of PSNR values between each RGB component obtained in our previous experiments does not affect depth perception.

6. CONCLUSION

In this work, we reported several formats available for stereoscopic representation and visualization, each one designed for a specific device or system. From that, we observed that the compression methods available for 2D videos could also be used by stereoscopic videos, but that could affect depth perception depending on the type of visualization technique used, highlighting the lack of a standard coding method exclusive for stereoscopic videos that would be generic and compatible among different visualization systems. We then showed that the anaglyphic is a stereoscopic visualization method that is simple to implement, does not require expensive or complex equipment to be visualized and can reduce data from a stereo video by half. Only if we could reverse it to its original stereo pair, the anaglyphic method would be a potential candidate for a generic stereoscopic compression process. Therefore, we proposed a technique for anaglyph reversal, with the creation of a Color Index Table that stores data from the color components discarded during the anaglyph process.

Our experiment showed that this reversion process is viable indeed. We were able to recreate a stereo pair of images from its respective green-magenta anaglyph with an average of 34,52 dB of PSNR and good quality when compared to the original one. The overhead of data with the addition of the Color Index Table in the anaglyph process was little, with an average reduction of 63.09% of the file size generated by it. Moreover, we found that the image resulted from the reversion process could still be transformed in anaglyph with no loss of depth perception.

The reversion technique involves converting images from RGB to YC_bC_r color spaces, which leads to rounding of floats values. That affected PSNR measurement. As a future work, we'll be restructuring our implementation in order to get better rounding values. We will also add more complexity in the technique to explore pixel displacement between the images of the stereo pair, in order to increase the quality of the reverted image.

7. REFERENCES

[1] Andrade, L., Dolosic, P., Goularte, R. 2010. *Construção de uma Base de Vídeos Estereoscópicos*. Technical Report. ICMC-University of São Paulo, São Paulo, Brazil. Available at http://www.icmc.usp.br/~biblio/BIBLIOTECA/rel_tec/RT_351.pdf.

[2] Andrade, L. A., Goularte, R. 2009. Anaglyphic stereoscopic perception on lossy compressed digital videos. In *Proceedings of the XV Brazilian Symposium on Multimedia and the Web* (WebMedia '09). Fortaleza, v.1, n.1, 226-233. DOI=http://doi.acm.org/10.1145/1858477.1858506.

[3] Andrade, L. A., Goularte, R. 2010. Uma Análise da Influência da Subamostragem de Crominância em Vídeos Estereoscópicos Anaglíficos. In *Proceedings of the XVI Brazilian Symposium on Multimedia and the Web* (WebMedia '10).

[4] Azevedo, E., Conci, A. 2003. *Computação gráfica: teoria e prática*. Campus, Elsevier, Brazil.

[5] Bradski, G., Kaehler, A. 2008. *Leaning OpenCV*. O'Reilly, United States.

[6] Dubois, E. 2001. A projection method to generate anaglyph stereo images. *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on* (May, 2001), 1661-1664. DOI= http://dx.doi.org/10.1109/ICASSP.2001.941256.

[7] Fehn, C., Kauff, P., Op de Beeck, M., Ernst, F., IJsselsteijn, W., Pollefeys, M., Van Gool, L., Ofek, E., Sexton, I. 2002. An Evolutionary and Optimised Approach on 3D-TV. In

Proceedings of International Broadcast Conference, 357-365.

[8] LG autostereoscopic mobile phone, available at http://www.lg.com/uk/mobile-phones/all-lg-phones/LG-android-mobile-phone-P920.jsp.

[9] Lipton, L.1982. *Foundations of the Stereoscopic Cinema: a study in depth*. Van Nostrand Reinhold Company Inc.

[10] Lipton, L. 1997. Stereo-Vision Formats for Video and Computer Graphics. *Proc. SPIE*, 3012, 239 (February. 1997). DOI=http://dx.doi.org/10.1117/12.274462.

[11] Mendiburu, B. 2009. *3D Movie Making Stereoscopic Digital Cinema from Script to Screen*. Focal Press, Elsevier, Brazil.

[12] Nintendo 3DS, available at http://www.nintendo.com/3ds/hardware.

[13] Richardson, I. E. 2003. *H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia*. Wiley, England.

[14] Smolic, A.; Mueller, K.; Merkle, P.; Kauff, P.; Wiegand, T. 2009. 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, 2009* (May, 2009), 1-4, 6-8. DOI= http://dx.doi.org/10.1109/PCS.2009.5167358.

[15] Sony 3D TV Technology, available at http://www.sony.net/united/3D/#technology/3dtv/.

[16] StereoGraphics Corporation. 1997. *Stereographics® Developers' Handbook: Background on Creating Images for CrystalEyes and SimulEyes*.

[17] Sumei, L., Chunping, H., Yicai, Y., Xiaowei, S., Lei, Y. 2009. Stereoscopic Video Compression Based on H.264 MVC. *Image and Signal Processing, 2009. CISP '09. 2nd International Congress on* (October, 2009), 1-5, 17-19. DOI= http://dx.doi.org/10.1109/CISP.2009.5301218.

[18] Tam, W. J., Zhang, L. 2006. 3D-TV Content Generation: 2D-to-3D Conversion. *Multimedia and Expo, IEEE International Conference on* (July 2006), 1869-1872. DOI= http://doi.ieeecomputersociety.org/10.1109/ICME.2006.262919.

[19] Vetro, A. 2010. Representation and Coding Formats for Stereo and Multiview Video. In Studies in Computational Intelligence. Springer Berlin / Heidelberg, 51-73. DOI= http://dx.doi.org/10.1007/978-3-642-11686-5_2.

[20] Winkler, S. 2005. *Digital Video Quality: vision model and metrics*. Wiley, England.