RevGlyph - A Technique for Reverting Anaglyph Stereoscopic Videos

ABSTRACT

In order to visualize stereoscopic videos, it is necessary a pair of videos of a same scene horizontally displaced – the stereo pair –, thus requiring twice the space to be stored. The anaglyph conversion is a technique in which color components of the stereo pair are removed, being the remaining color components joined to form a single video. With that, it's possible to reduce at least by the half the amount of data to be stored. Unfortunately, the anaglyphic video is not supported by other stereoscopic visualization methods – they require the stereo pair. Our proposal is to create an anaglyph reversion technique that would enable to recreate the stereo pair from an anaglyph video, which could make stereo information available for other visualization methods. Such reversion is not straightforward, since during the anaglyph conversion some color data is lost. In this paper, we introduce RevGlyph, an approach to the anaglyph reversion by storing the removed color components in a special structure that we call "Color Index Table". We then discuss the results obtained from our experiments with this technique, achieving a good compression rate of 79.6395% compared to the original image. The results also showed that the process doesn't interfere with the stereo depth perception.

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, Experimentation, Standardization.

Keywords

Stereoscopic visualization, Stereoscopic video coding, anaglyph video.

1. INTRODUCTION

Stereoscopic videos, commonly known as 3D videos, are formed by a pair of videos – called the 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 [11]. 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[1], cell phones[2] and video games[3], 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.

Data compression is a constant need for video storage and transmission, especially nowadays with professional and homemade videos being recorded in high definitions and transmitted massively over the internet on sites like Youtube[4] and NetFlix[5]. With stereoscopic videos, this need for compression is increased, since it is necessary to store the two videos that form the stereo pair. The scenario can be even worse with new technologies like free viewpoint video (FVV) [12] in which several video streams are reproduced giving the viewer the possibility of visualizing the same scene from different angles. With this constant increase of data to be stored and transmitted, new techniques for data compression are needed.

Revising the literature, one can notice that there's no exclusive coding specially designed for stereoscopic videos. The available strategies can be divided into two groups: Lipton's Method [10], and methods involving video and depth [14]. Lipton's Method describes formats for stereoscopic video representation, being the stereo pair stored in a single video container with double of data than a regular 2D video stream. Methods involving video and depth were created in order to meet the demand of future technologies like FVV and autostereoscopic displays. They use new concepts like depth maps [14] and image correlation [1] seeking for better ways of representing and storing this type of media.

Regarding Lipton's Method, since the stereo pair is kept, it can be used by any stereoscopic visualization system – it is a generic approach, independent of the visualization method. However, keeping the stereo pair results in double the video storage needs. Methods involving video and depth, on the other hand, may not require the stereo pair – some techniques store only one video stream with a depth map attached. This depth map is used to render the second view. These methods also use well-known compression techniques like MPEG-2 and H.264 with some adaptations to explore redundancies that appear on stereoscopic videos. However, there are several variations and formats of such methods, each developed to attend one or another method of stereoscopic visualization – anaglyph, polarized light, shutter

[1] http://blog.us.playstation.com/2011/06/06/new-3d-display-and-ps3-accessories-debut-at-e3/

[2] http://www.lg.com/uk/mobile-phones/all-lg-phones/LG-android-mobile-phone-P920.jsp

[3] http://www.nintendo.com/3ds/hardware

[4] http://www.youtube.com/

[5] http://www.netflix.com/

glasses or autostereoscopic – which brings incompatibility among different systems [14]. Moreover, since lossy compression is used, some data loss will occur in color information [5], which may compromise depth perception for some visualization systems. With that said, it's possible to realize that there's a lack of a generic method specially designed for stereoscopic video coding, compatible with different types of 3D visualization, providing good quality, no loss of depth perception and good compression rate.

One approach to tackle the problem of stereoscopic video compression is to convert the stereo pair into a single anaglyph video stream. The anaglyph method operates by merging the stereo pair into a single video, which, in turn, may be compressed using lossless techniques. As a consequence, we obtain a higher compression rate. However, the anaglyph method is compatible only with the anaglyph visualization method. In order to create generic stereoscopic coding methods, compatible with all kinds of stereoscopic visualization, the reversion of the anaglyph into the original stereo pair becomes a need. Such reversion technique is not trivial though, since the anaglyph conversion causes loss of color information, and it must be retrieved somehow.

This paper demonstrates that this reversion is possible and proposes a technique – RevGlyph – based on a special structure that we call "Color Index Table", in which the color components removed during anaglyph conversion are subsampled and stored. With this technique, we were able to achieve 79.6395% of compression and to create a reversed stereo pair with no loss in depth perception. Since the stereo pair is a need for all stereoscopic visualization methods, this technique becomes a generic approach for stereoscopic video coding. Moreover, the reverted stereo pair may be coded by any other state of the art coding technique – the proposed technique does not exclude a joint use with other available ones.

This paper is organized in 6 sections, including this Introduction (Section 1). Section 2 reviews related work regarding stereoscopic video coding. Section 3 presents the RevGlyph technique. Section 4 has details over the experiments taken, with Section 5 analyzing the results obtained. Section 6 concludes this paper with the main achievements found and key points for future research.

2. RELATED WORK

The MPEG group has well-known standard methods for video coding, even with extensions for multiple streams of videos, like MPEG-2 Multiview Profile [14]. However, there is no standard coding 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. The stereo pair can be stored in a single video file in two formats: above-and-below or side-by-side. They differ in the way the two images of the stereo pair are organized. In the above-and-below format, image from the left is placed at the top of a frame and image from the right is placed at the bottom. In the side-by-side format, image from the left is placed at the left of the frame and image from the right is placed at the right. As the stereo pair is kept without any modification, Lipton's Method is independent from device, which means that any stereoscopic visualization system is capable of reproducing it.

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 – which they call views –, order of complexity and types of data involved, resulting in six classes: Conventional Stereo Video (CSV), Multiview Video Coding (MVC), Video plus depth (V+D), Multiview plus depth (MVD), Layered Depth Video (LDV) and Depth Enhanced Stereo (DES).

CSV and MVC are similar to Lipton's Method, differing on the number of views each one holds. CSV handles only a pair of video signals – the stereo pair. MVC is an extension to when more than two views are used. A pair of adjacent views will form a stereo pair, enabling the viewer to see a scene from different angles. H.264 has an extension to include MVC [14] and some authors like [16] study ways to improve this method for better performance.

The next four formats are strategies involving video and depth. V+D is a format 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, called virtual views. MVD is a combination of MVC and V+D properties, which means that multiple views and multiple depth maps are sent, increasing the number of virtual views that can be created. LDV is an alternative for MVD, in which a set of additional layers are stored together with the video signal and depth map to generate multiple virtual views, thus reducing the number of video signals to be stored. 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 depth maps and layers, providing compatibility among different formats, since each one uses only the types of data needed. Even though it becomes generic, DES brings the problem of storing additional information that may or may not be used, depending on the device that the video will be played.

Notice that either in Lipton's Method or methods involving video and depth, the amount of data needed is far more when compared to a single 2D video stream. New researches were made trying to find new ways of achieving more compression from stereoscopic videos. Some authors explored a new type of redundancy that is present in stereoscopic media: the worldline correlation [1][2]. Since images from a stereo pair have lots of correspondent pixels horizontally displaced, one of the images can be used as basis for predicting the other one, similar to the motion estimation performed to remove temporal redundancies from videos.

Vetro [17] 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).

Through the review of related work, we observe 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; and b) achieve high compression rates preserving image quality and depth perception, while keeping the visualization independence. In the following sections, we present an anaglyph reversion technique that enables a convenient compression of stereoscopic videos by storing them in the anaglyph format – tackling problem b), which may be in turn reversed for the usage of other visualization methods that don't support the anaglyph – tackling problem a).

3. THE REVGLYPH TECHNIQUE

In order to present the anaglyph reversion technique, first we need to understand how the anaglyph conversion works and preprocessing steps needed to enable the reversion.

The anaglyph method is an important low cost and simple way of coding a stereo pair to be visualized using proper glasses. In this method, we remove some color components from each video, merging the remaining components into a third single video. Afterwards, by using glasses with lenses that mimics the color components removed, thus acting as filters, the color channels are separated again and each eye sees different stereo images, leading to the retinal disparity[6] and resulting in the stereoscopic effect.

Figure 1 illustrates the anaglyph conversion. The stereo pair is composed of a series of frames, being each frame in turn composed by a pair of images, one for the right eye and one for the left eye. To create an anaglyph, RGB channels of each frame are combined in a way to have information from both images [9]. Let a frame be formed by $R_1G_1B_1$ (the right eye image) and $R_2G_2B_2$ (the left eye image). The conversion takes the green channel (G_1) from the image for the right eye and channels R_2 and B_2 from the image for the left eye, 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 anaglyph method illustrated in Figure 1 is known as green-magenta. Other possible methods are: red-cyan ($R_1G_2B_2$) and blue-yellow ($R_2G_2B_1$).

Dubois [4] reported that the green-magenta method presents better results for anaglyphic visualization than the others. Therefore, in this work, we use the green-magenta color combination.

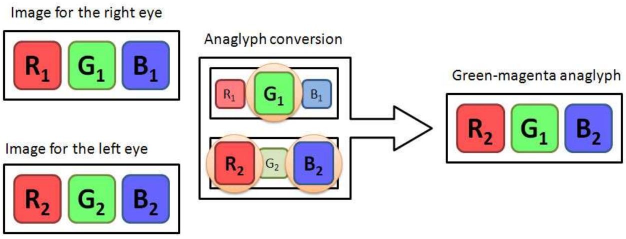

Figure 1 - Green-magenta anaglyph creation.

[6] The retinal disparity is the difference in distance between the positions of the image formed on each retina [15]

3.1 Creating the Color Index Table

A first attempt to recover the stereo pair from an anaglyph video is to store the discarded color information into a special data structure that we call "Color Index Table". Following Figure 1 example, this table will be formed by color information from channels R_1, G_2 and B_1 and stored together with the $R_2G_1B_2$ video file. 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 we are only interested in color information to build the Color Index Table, a better approach is to use a color space conversion from RGB to YC_bC_r[7]. With that, 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 subsampled, reducing even more the amount of data needed to be stored in the table.

There are some possible chrominance subsampling combinations, presenting different tradeoffs between compression and color fidelity [13]. The 4:4:4 combination is the best in quality, but the worst in terms of information reduction, since it keeps the original data. The 4:1:1 combination is exactly the opposite, greatly reducing information but resulting in images with poor quality. Combinations 4:2:2 and 4:4:0 work by picking only two samples of C_b and C_r components for each four samples of Y component, thus reducing the chrominance resolution by half, still resulting in images with good quality. A vertical reduction is applied on 4:2:2 whereas a horizontal reduction is applied on 4:4:0 [8][13]. In this work, we'll use both 4:2:2 and 4:4:0 combinations and compare their results, in order to verify which one presents better results.

The Color Index Table is built by following 4 steps, depicted in Figure 2: (I) creation of a green-magenta anaglyph image ($R_2G_1B_2$) from the uncompressed stereo pair; (II) creation of another anaglyph ($R_1G_2B_1$) with the remaining color components, which we call "complementary anaglyph"; (III) conversion of both anaglyphs from RGB to YC_bC_r color space using either 4:2:2 or 4:4:0 subsampling; and (IV) discard of Y' component and storage of C_b' and C_r' components from the complementary anaglyph into the color index table. Afterwards, lossless compression is applied to both green-magenta anaglyph and the color index table, resulting in a single compressed file.

3.2 Anaglyph Reversion

The anaglyph reversion consists of recreating the stereo pair using the color index table and the green-magenta anaglyph, 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 Y component and applying the conversion from YC_bC_r to RGB to recreate the complementary anaglyph. Since the Y component from the complementary anaglyph was discarded during the conversion, we make use of the one from the green-magenta anaglyph. Before the YC_bC_r to RGB conversion, we need to adjust the color resolution by duplicating each chrominance component to every 4 samples of luminance component, thus returning to 4:4:4 combination.

[7] There is a standard equation for this conversion [6]

Now that we have both the green-magenta and the complementary anaglyph, we just need to reorganize the color components belonging 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 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.

4. EXPERIMENTS

For our experiments, we have implemented the conversion and reversion technique in C++ with the aid of the OpenCV library[8]. We have then applied the technique to a set of 32 stereo images, extracted from a test database of stereoscopic videos[9], being the frame extracted from each video the first in which we perceived the stereoscopic effect clearly. Experiments ocurred in two batches: the first one using 4:2:2 subsampling during anaglyphic conversion phase and the other one using 4:4:0 subsampling.

Different from 2D videos, standard stereo video data sets are coarsely found in literature. That was the reason to bulid our own video data set, in which the images were classified and grouped following ITU-T recommended criteria [7]: images have to present differences on brightness, saturation and contrast. Brightness deals 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. Since we're dealing with anaglyph images, color is an important aspect to analyze. 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. Since these criteria are not mutually exclusive, one or more of them may appear in the same image. Therefore, some images were replicated during the tests.

The tests were based on two aspects: the analysis of the compression rate obtained after the anaglyph convertion and on the Peak Signal-to-Noise Ratio (PSNR) [18] of the reverted stereo pair in comparison with the original one. We have calculated the compression rate obtained both with and without the color index table, in order to analyze how much data our technique would add as an overhead to the compressed image. PSNR was calculated using a free version of the MSU Video Quality Measurement Tool[10] (MSU VQMT). For each criteria, we have took the PSNR of each color component in the RGB color space and then calculated the average of them.

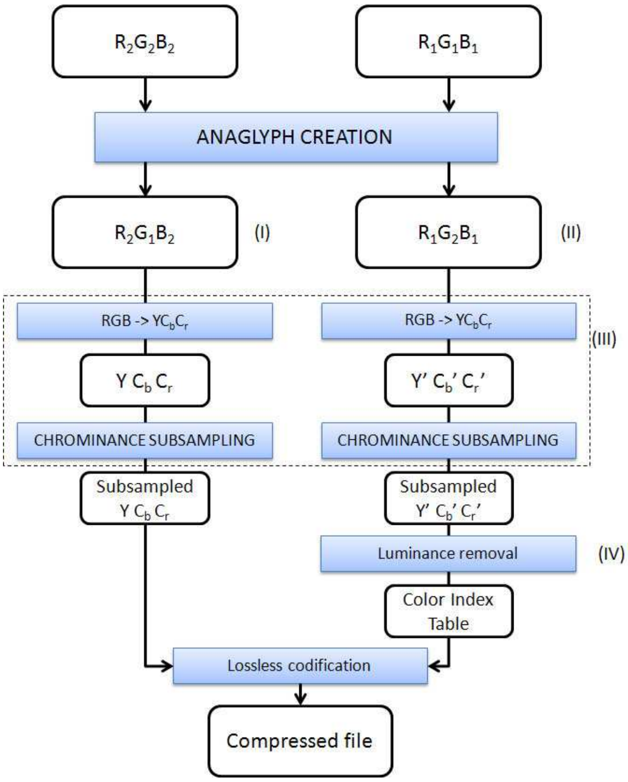

Figure 2 - Anaglyph conversion process using the color index table.

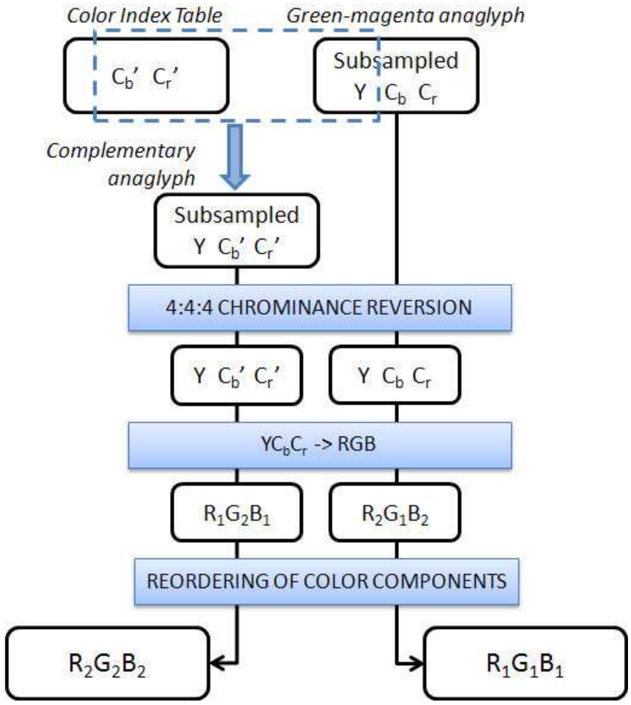

Figure 3 - Anaglyph reversion technique, using the color index table and luminance component from green-magenta anaglyph to recreate the complementary anaglyph.

---

[8] This library has several implementations of image processing functions [3] and can be downloaded at http://sourceforge.net/projects/opencvlibrary/

[9] The database was created by our group and is available online. URL for database was ommitted in this first version of the paper for the sake of the double-blind review process

[10] http://compression.ru/video/quality_measure/vqmt_download_en.html

5. RESULTS AND DISCUSSION

Table 1 shows the average PSNR obtained for each of the three criteria defined, by using 4:2:2 and 4:4:0 subsampling combinations. The last column presents the overall results – excluding duplicated images that appeared on more than one criterion.

Table 1 - Average PSNR values obtained by the anaglyph reversion technique using 4:2:2 and 4:4:0 subsampling combinations

	PSNR Brightness (dB)	PSNR Saturation (dB)	PSNR Contrast (dB)	PSNR Overall (dB)
4:2:2	40.4639	40.6885	41.6161	40.8888
4:4:0	40.6428	40.8900	41.3111	40.9054
Difference	0.1789	0.2015	0.2851	0.0166

As expected, from the results, we can observe that the PSNR values were very similar both between the three criteria and between the subsampling combinations as well. The maximum PSNR difference between the three criteria was of 1.1522 dB and the maximum PSNR difference between 4:2:2 and 4:4:0 subsampling combinations was of 0.2851 dB, being of only 0.0166 dB in the overall results. From this, we conclude that the differences in the stereo pair (right eye and left eye points of view) cause no meaningful impact in the final image quality

Table 2 shows the average compression rate obtained by using 4:2:2 and 4:4:0 subsampling combinations. First column presents the compression rate when the color index table is not present, whereas the second column presents the compression rate including the color index table, with the third column showing the difference of values – the data overheard of the color index table.

Table 2 - Average compression rate obtained by the anaglyph reversion technique using 4:2:2 and 4:4:0 subsampling combinations

	Compression rate without the color index table (%)	Compression rate with the color index table (%)	Data Overhead (%)
4:2:2	84.6936	76.6849	8.0087
4:4:0	86.1702	79.6395	6.5307
Difference	1.4766	2.9546	1.4780

Considering all these results, the objective analysis performed showed that 4:4:0 subsampling combination is superior. Since the depth in stereo video deals with perception, a subjective analysis (ITU-T [7]) may reveal a scale between the 4:2:2 and 4:4:0 combinations. A comprehensive analysis like that was left as future work. A preliminary analysis made by three individuals, qualitatively comparing reversed images generated using both combinations, showed that they presented similar quality. Moreover, there was no loss in depth perception. Therefore, the RevGlyph technique uses 4:4:0 combination.

Regarding the data overhead needed for the storage of the color index table, one can observe that it is very low – 6.53% using 4:4:0 combination – in face of the advantage that it is now possible to perform the anaglyph reversion.

In Sections 5.1 and 5.2 we discuss two points regarding the anaglyph reversing technique that will guide us towards future research.

5.1 Image Quality and Depth Perception

Figure 4 shows one of the reverted images obtained by using the RevGlyph technique with 4:4:0 subsampling. By observing the reverted stereo pair, it's clearly perceptible the presence of crosstalk[11] on the top part of the train and in the woods of the image from the left. This is a direct consequence of using the luminance component from the green-magenta anaglyph to reconstruct the complementary one.

Even though crosstalk is present, the resulted image presents good visual quality, with the objects from the scene being clearly discernible. Regarding depth perception, another experiment was made: we took the 32 reverted images and transformed them into green-magenta anaglyphs. Figure 5 shows a sample: 5(A) shows the anaglyph from the original stereo pair and 5 (B) shows the anaglyph resulted from this transformation. When using appropriated glasses (Roscolux #139: Primary Green and #339: Broadway Pink color filters), one can notice that both images are very similar, and depth perception is not lost. That ensures that the reverted stereo pair does not lose depth characteristics when using the RevGlyph technique.

5.2 PSNR Analysis

During the experiments, the best PSNR achieved using 4:4:0 combination was of 46,8840 dB. This is a low value when compared to the PSNR range of 0 through 100. There are two factors involved in this result: the subsampling and the crosstalk.

The chrominance subsampling is a lossy process. When using either 4:2:2 or 4:4:0 combinations, chrominance resolution is reduced by a half. During the anaglyph reversion, the stereo pair returns to the 4:4:4 combination by duplicating chrominance values in the pixels that were not sampled. Such pixel values are not exactly the same from the original, thus leading to a lower PSNR value.

We took an experiment to verify how much the chrominance subsampling impacts on the PSNR value. In this experiment, the 32 images were just converted to YC_bC_r and subsampled using 4:4:0 combination, being afterwards reverted to 4:4:4 and then to RGB color space. The PSNR stayed around 50dB. Therefore, it can be concluded that the RevGlyph technique introduces about 10dB of PSNR loss.

[11] Woods [19] presented on his paper several definitions regarding crosstalk. In this article, we are accepting the definition of crosstalk as being an "incomplete isolation of the left and right image channels so that one image leaks into the other"

Figure 4 - Stereo Pair created using the anaglyph reversion technique.

As discussed in Section 5.1, there's a presence of crosstalk in the reverted stereo pair. In other words, pixels appear displaced in several points of both images, which affect the pixel-by-pixel comparison of the PSNR measurement. That is because the Y component used to obtain the complementary anaglyph came from an anaglyph (the green-magenta one). During the reversion process, color components are rearranged causing the appearance of the crosstalk. However, if one transforms the reversed image into an anaglyph, these color components are rearranged again in a way that they are merged, which results in an image with no perceptible crosstalk and no depth perception as well.

(A)

(B)

Figure 5 - Comparison with green-magenta anaglyphs. (A) was created from the original stereo pair and (B) was created from the reverted one.

6. CONCLUSION

In this work, we discussed about the high amount of data to be processed and stored when dealing with stereoscopic video, which creates the need of new compression techniques. We also 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 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 at least the 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 the RevGlyph technique for anaglyph reversion, involving the creation of a green-magenta anaglyph and a data structure that we called "Color Index Table" that would store subsampled chrominance data discarded during the anaglyph process. Using the color index table and the luminance component from the green-magenta anaglyph, it's possible to restore the missing color components, resulting in the recreation of the stereo pair.

Our experiments were based on two implementations of the anaglyph reversion, one using 4:2:2 chrominance subsampling combination and the other using 4:4:0. Both combinations revealed the viability of this reversion process, presenting similar results, with a littler advantage for the 4:4:0 combination in terms of PSNR and compression rate. We were able to recreate a stereo pair of images from its respective green-magenta anaglyph with an average PSNR of 40.9054 dB and good quality when compared to the original one. The compression rate obtained was of 79.6395%, being 6.5307% the data overhead from the color index table. Moreover, we showed that the image resulted from the reversion process could still be transformed into another green-magenta anaglyph with no loss of depth perception.

More research is needed in order to find ways to eliminate or alleviate the crosstalk. A possible solution we'll be looking into future work is the integration of worldline correlation to the reversion process, that is, the creation of a structure that holds the position for every similar point in both images that appears in different horizontal positions due to the displacement from camera capture. With that information, we would be able to reorganize values from the luminance component of the green-magenta anaglyph to reconstruct a similar luminance component for the complementary anaglyph. This will directly impact not only on an image with greater visual quality, but also on better PSNR results.

7. REFERENCES

[1] Adikari, A.B.B.; Fernando, W.A.C.; Arachchi, H.K.; Loo, K. 2005. A H.264 compliant stereoscopic video codec. *Electrical and Computer Engineering, 2005. Canadian Conference on* (May 2005), 1614-1617, 1-4. DOI=10.1109/CCECE.2005.1557292

[2] Balasubramaniyam, B; Edirisinghe, E; Bez, H. An extended H.264 CODEC for stereoscopic video coding, *Proc. SPIE* (January, 2005), 5664, 116. DOI=10.1117/12.587583

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

[4] 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= 10.1109/ICASSP.2001.941256

[5] Gonzalez, R. C.; Woods, R. E. 2008. *Digital Image Processing*. 3rd ed. Pearson Prentice-Hall, United States.

[6] ITU-R. 2011. *Recommendation ITU-R BT.601-7: Studio encoding parameters of digital television for standard 4:3*. International Telecommunication Union.

[7] ITU-T. 2004. *Objective perceptual assessment of video quality: Full reference television.* International Telecommunication Union.

[8] KERR, D. A. 2009. Chrominance Subsampling in Digital Images. Available at <http://dougkerr.net/pumpkin/articles/Subsampling.pdf>. Last access September 7th 2011.

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

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

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

[12] Merkle, P.; Smolic, A.; Muller, K.; Wiegand, T. 2007. Efficient Prediction Structures for Multiview Video Coding. *Circuits and Systems for Video Technology, IEEE Transactions on* (November, 2007), 1461-1473. DOI=10.1109/TCSVT.2007.903665

[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=10.1109/PCS.2009.5167358

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

[16] Sumei Li; Chunping Hou; Yicai Ying; Xiaowei Song; Lei Yang. 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=10.1109/CISP.2009.5301218

[17] Vetro, A. 2010. Representation and Coding Formats for Stereo and Multiview Video. *Studies in Computational Intelligence*, 51-73. DOI=10.1007/978-3-642-11686-5_2

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

[19] Woods, A. J. 2011. How are crosstalk and ghosting defined in the stereoscopic literature?. *Proc. SPIE* (January, 2011), 7863, 78630Z. DOI=10.1117/12.877045