A PROJECTION METHOD TO GENERATE ANAGLYPH STEREO IMAGES

Eric Dubois

School of Information Technology and Engineering
University of Ottawa, Ottawa, Ontario, Canada K1N 6N5
edubois@uottawa.ca

ABSTRACT

An anaglyph image allows the perception of depth when observed through colored glasses such as the familiar red/blue glasses. Although the method is very old, the techniques used to generate anaglyph images are very empirical. This paper describes a projection method to generate anaglyph stereoscopic images using the spectral absorption curves of the glasses, the spectral density functions of the display primaries and the colorimentric properties of the human observer.

1. INTRODUCTION

This paper describes a method to form an anaglyph stereo image from the left and right color images of a stereo pair. The anaglyph method was patented in 1891 by Louis Ducos du Hauron, but similar methods had been demonstrated previously by W. Rollmann in 1853 and J.C. D'Almeida in 1858 [1]. The word *anaglyph* is from the Greek $\alpha\nu\alpha$ = again, $\gamma\lambda\mu\phi\eta$ = sculpture. In the classic method, used for monochrome stereo images, the left view in blue (or green) is superimposed on the same image with the right view in red. When viewed through spectacles of corresponding colors but reversed, the three-dimensional effect is perceived.

Color anaglyphs have also been proposed; a method to produce a color anaglyph using Photoshop has been described in [2] for red/cyan glasses. Although other stereoscopic visualization methods such as those using polarized or shuttered glasses can give better results, the anaglyph method is the only way that stereoscopic images can be viewed on ordinary television sets or computer screens with no special hardware other than inexpensive colored glasses.

There is very little literature on the production of anaglyph images, and what exists is very empirical. The method proposed in this paper is adapted to the spectral absorption curves of the left and right filters of the anaglyph glasses. A projection technique is used to compute the anaglyph image that yields an image pair (after the glasses) as close as possible to the desired stereo pair.

2. FORMULATION

2.1. The input stereo pair

The input to the process is a true-color stereo pair $V'_{lj}(x)$ and $V'_{rj}(x)$, $j = 1, 2, 3$ and $x \in \mathcal{L}$. It is assumed that the three components $j = 1, 2, 3$ are gamma-corrected RGB (in that order) that can be directly displayed on a standard CRT monitor; the 'prime' symbol denotes gamma-corrected signals. $\mathcal{L}$ is the sampling raster for the image which is arbitrary and can be either spatial or spatiotemporal. In an ideal stereoscopic system, the three components go through the display gamma, denoted by the function $g(\cdot)$, and excite the display RGB phosphors. The spectral density functions of the RGB display phosphors are denoted $d_j(\lambda)$, $j = 1, 2, 3$ respectively. Thus, the spectral density of the light emanating from point x in the left and right images is given by

$$Q_l(\lambda, x) = \sum_{j=1}^{3} V_{lj}(x)d_j(\lambda), \qquad (1)$$

$$Q_r(\lambda, x) = \sum_{j=1}^{3} V_{rj}(x)d_j(\lambda) \qquad (2)$$

where $V_{lj}(x) = g(V'_{lj}(x))$ and $V_{rj}(x) = g(V'_{rj}(x))$. The color perceived at point x in the left and right images is determined by the projection of $Q_l(\lambda, x)$ and $Q_r(\lambda, x)$ onto the visual subspace using color-matching functions $\bar{p}_k(\lambda)$ for the chosen set of primaries:

$$\begin{aligned}
\widetilde{V}_{lk}(x) &= \int Q_l(\lambda, x)\bar{p}_k(\lambda)\, d\lambda \\
&= \sum_{j=1}^{3} V_{lj}(x) \int \bar{p}_k(\lambda)d_j(\lambda)\, d\lambda \\
&= \sum_{j=1}^{3} c_{kj}V_{lj}(x), \quad k = 1, 2, 3.
\end{aligned} \qquad (3)$$

The integral is over the wavelengths of the visible spectrum, approximately 370 nm to 730 nm. Thus, in matrix notation,

$\widetilde{V}_l(x) = CV_l(x)$, where

$$[C]_{kj} = c_{kj} = \int \bar{p}_k(\lambda)d_j(\lambda)\,d\lambda, \qquad (4)$$

$V_l(x) = [V_{l1}(x)\ V_{l2}(x)\ V_{l3}(x)]^T$ and $\widetilde{V}_l(x) = [\widetilde{V}_{l1}(x)\ \widetilde{V}_{l2}(x)\ \widetilde{V}_{l3}(x)]^T$. Similarly, $\widetilde{V}_r(x) = CV_r(x)$. In an ideal stereoscopic visualization system, the left eye sees only the image defined by $\widetilde{V}_l(x)$ and the right eye sees only the image defined by $\widetilde{V}_r(x)$.

We use the XYZ coordinate system, and so define the color-matching functions $\bar{p}_k(\lambda)$, $k = 1, 2, 3$ to be the standard $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ respectively. These functions are tabulated and graphed in [3].

The display phosphor spectral densities for a Sony Trinitron monitor are shown in Fig. 7.16 of [4] (the numerical values were provided by Charles Poynton). Using these spectral densities, the matrix C is

$$C = \begin{bmatrix} 0.4641 & 0.3055 & 0.1808 \\ 0.2597 & 0.6592 & 0.0811 \\ 0.0357 & 0.1421 & 0.9109 \end{bmatrix} \qquad (5)$$

This matrix is similar to standard ones for converting from various RGB spaces to XYZ, e.g., the Rec. 709 matrix on page 148 of [4], but is slightly different. If different phosphors are used, a different matrix would result.

The value of the stereoscopic image at each point x can be considered to be an element of a six-dimensional vector space S_6. Arranged as a column matrix, we have

$$\widetilde{V}(x) = \left[\widetilde{V}_{l1}(x)\ \widetilde{V}_{l2}(x)\ \widetilde{V}_{l3}(x)\ \widetilde{V}_{r1}(x)\ \widetilde{V}_{r2}(x)\ \widetilde{V}_{r3}(x)\right]^T. \qquad (6)$$

We can form a basis for this space using the columns of C as follows:

$$c_{li} = \begin{bmatrix} c_{1i} & c_{2i} & c_{3i} & 0 & 0 & 0 \end{bmatrix}^T$$
$$c_{ri} = \begin{bmatrix} 0 & 0 & 0 & c_{1i} & c_{2i} & c_{3i} \end{bmatrix}^T \qquad (7)$$
$$i = 1, 2, 3$$

In terms of this basis, we have

$$\widetilde{V}(x) = \sum_{j=1}^{3} V_{lj}(x)c_{lj} + \sum_{j=1}^{3} V_{rj}(x)c_{rj}. \qquad (8)$$

If we define the 6×6 matrix

$$C_2 = \begin{bmatrix} c_{l1} & c_{l2} & c_{l3} & c_{r1} & c_{r2} & c_{r3} \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix} \qquad (9)$$

then we can write in matrix form

$$\widetilde{V}(x) = C_2 V(x). \qquad (10)$$

The set of realizable stereoscopic images has values that lie in the convex subset of S_6

$$\left\{ \sum_{j=1}^{3} v_{lj}c_{lj} + \sum_{j=1}^{3} v_{rj}c_{rj} \mid 0 \le v_{lj} \le 1, 0 \le v_{rj} \le 1, \right.$$
$$\left. j = 1, 2, 3 \right\}. \qquad (11)$$

2.2. Visualization of an anaglyph image

A single anaglyph image is denoted $V'_{aj}(x)$, $j = 1, 2, 3$, $x \in \mathcal{L}$ where the three components are in the same gamma-corrected RGB display primary system as the stereo pair of the previous section. Thus, the spectral density of the light emitted from the screen at x is given by

$$Q_a(\lambda, x) = \sum_{j=1}^{3} V_{aj}(x)d_j(\lambda) \qquad (12)$$

where $V_{aj}(x) = g(V'_{aj}(x))$.

The light from the CRT passes through two filters with spectral absorption functions $f_l(\lambda)$ and $f_r(\lambda)$ before arriving at the left and right eyes respectively. Thus the light spectral distribution at the left and right eyes is $Q_a(\lambda, x)f_l(\lambda)$ and $Q_a(\lambda, x)f_r(\lambda)$ respectively. The corresponding sets of XYZ tristimulus values are

$$\widetilde{U}_{lk}(x) = \int Q(\lambda, x)f_l(\lambda)\bar{p}_k(\lambda)\,d\lambda$$
$$= \sum_{j=1}^{3} V_{aj}(x) \int \bar{p}_k(\lambda)d_j(\lambda)f_l(\lambda)\,d\lambda \qquad (13)$$
$$= \sum_{j=1}^{3} a_{lkj}V_{aj}(x)$$

or in matrix form $\widetilde{U}_l(x) = A_l V_a(x)$, where

$$[A_l]_{kj} = a_{lkj} = \int \bar{p}_k(\lambda)d_j(\lambda)f_l(\lambda)\,d\lambda. \qquad (14)$$

Similarly, $\widetilde{U}_r(x) = A_r V_a(x)$, where

$$[A_r]_{kj} = a_{rkj} = \int \bar{p}_k(\lambda)d_j(\lambda)f_r(\lambda)\,d\lambda. \qquad (15)$$

The goal is for the stereo pair perceived by viewing $\widetilde{U}_l(x)$ and $\widetilde{U}_r(x)$, $x \in \mathcal{L}$, to be as similar as possible to the ideal one perceived by viewing $\widetilde{V}_l(x)$ and $\widetilde{V}_r(x)$, $x \in \mathcal{L}$.

Of course, we can't make $\widetilde{U}_l(x) = \widetilde{V}_l(x)$ and $\widetilde{U}_r(x) = \widetilde{V}_r(x)$ in general, since the filters $f_l(\lambda)$ and $f_r(\lambda)$ each block certain wavelength bands. Specifically, if we want

to reproduce a feature that is dark in the left view and bright in the right view due to disparity, the light emitted at point x must lie mostly in the stopband of the left filter and in the passband of the right filter. Thus, the two filters must be complementary in some way. Fig. 1 shows the transmission characteristic for two Roscolux filters that are similar to the red and blue of commercial anaglyph glasses. Assuming that the red filter is on the left and that the blue filter is on the right, the matrices A_l and A_r corresponding to these filters are given by

$$A_l = \begin{bmatrix} 0.3185 & 0.0769 & 0.0109 \\ 0.1501 & 0.0767 & 0.0056 \\ 0.0007 & 0.0020 & 0.0156 \end{bmatrix} \quad (16)$$

$$A_r = \begin{bmatrix} 0.0174 & 0.0484 & 0.1402 \\ 0.0184 & 0.1807 & 0.0458 \\ 0.0286 & 0.0991 & 0.7662 \end{bmatrix} . \quad (17)$$
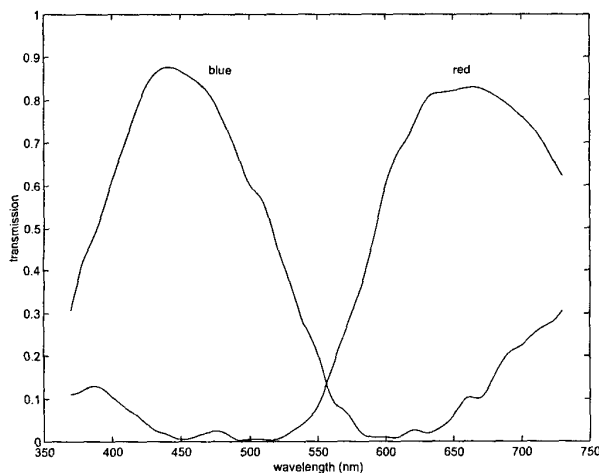

Fig. 1. Transmission as a function of wavelength of two Roscolux filters: orange red (#25) and brilliant blue (#69).

The stereoscopic image values formed by viewing the anaglyph image through spectacles with the colored filters also lie in the six-dimensional space S_6. However, they are constrained to lie in a three-dimensional subspace. Define the following three vectors in S_6:

$$r_j = \begin{bmatrix} a_{l1j} & a_{l2j} & a_{l3j} & a_{r1j} & a_{r2j} & a_{r3j} \end{bmatrix}^T , \quad j = 1, 2, 3. \quad (18)$$

Then

$$\widetilde{U}(x) = \sum_{j=1}^{3} V_{aj}(x) r_j \quad (19)$$

which lies in $\mathcal{R} = \text{span}(r_1, r_2, r_3)$, a three-dimensional subspace of S_6. The set of all *realizable* anaglyph stereoscopic images lies in the convex subset of $\mathcal{R}$

$$\left\{ \sum_{j=1}^{3} v_{aj} r_j \mid 0 \le v_{aj} \le 1, j = 1, 2, 3 \right\}. \quad (20)$$

If we define the matrix

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \end{bmatrix} = \begin{bmatrix} A_l \\ A_r \end{bmatrix} \quad (21)$$

then equation (19) can be expressed in matrix form as

$$\widetilde{U}(x) = R V_a(x). \quad (22)$$

3. FORMATION OF AN ANAGLYPH IMAGE

Given the formulation of the previous section, the formation of an anaglyph image can now be posed as an optimization problem. Given a stereoscopic pair $V_l(x)$, $V_r(x)$, $x \in \mathcal{L}$, we seek an anaglyph image $V_a(x)$, $x \in \mathcal{L}$ with $0 \le V_{aj}(x) \le 1$ such that the perceived image $\widetilde{U}$ is as similar as possible as the input image $\widetilde{V}$. Not only can we not obtain equality of $\widetilde{U}$ and $\widetilde{V}$ because of the spectral characteristics of the two filters, we must also take into account the reduction in luminance due to the overall attenuation of the filters.

In order to obtain the anaglyph image by optimization, we need an error metric that computes numerically the subjective difference between a stereo pair and an anaglyph type approximation. I know of no such error metric, so this initial study uses the following simplifications:

- The approximation is carried out independently at each sample location.

- The error metric at each point is a weighted squared error between $\widetilde{U}$ and $\widetilde{V}$.

- A global scaling of the V_{aj} is used to account for the attenuation of the filters.

Given these assumptions, the $V_{aj}(x)$ are determined by applying the projection theorem. The resulting values, which are not in the desired range of [0,1], are then scaled with fixed scaling factors s_j applied to all pixels of the image and then truncated to the [0,1] range. Experiments have shown that this approach can give a better three-dimensional effect than previously described empirical methods. It can easily be adapted to other filters and even to the shuttered anaglyph method.

In order to apply the projection theorem, we define an inner product on S_6 such that the resulting norm is a suitable

distance measure for our problem. A general inner product has the form

$$\langle v_1 \mid v_2 \rangle = v_1^T W v_2 \tag{23}$$

where W is a positive-definite matrix. The corresponding norm is

$$\|v\|^2 = \langle v \mid v \rangle = v^T W v. \tag{24}$$

If $W = I$, the 6×6 identity matrix, this results in a familiar Euclidean distance in the XYZ space (or more precisely in the Cartesian product of the XYZ space with itself). Use of other diagonal matrices W can allow weighting of the Y component more heavily than X and Z to favor reproduction of the correct luminance at the expense of greater color errors. Non-diagonal weighting matrices can correspond to distances with respect to other sets of primaries than XYZ. The projection approach is then to determine for each x the element of $\mathcal{R}$ that is closest in the sense of the chosen norm to $\widetilde{V}(x)$, i.e., find $\widehat{V}_a(x)$ such that $\|\widetilde{V}(x) - \sum_{j=1}^{3} \widehat{V}_{aj}(x) r_j\|$ is minimized. This is followed by scaling and clipping.

The method for finding the projection is standard; our approach is very similar to that described in Section 3.3.3 of [5]. We form the 3×3 Grammian matrix Φ,

$$\Phi = \begin{bmatrix} \langle r_1 \mid r_1 \rangle & \langle r_2 \mid r_1 \rangle & \langle r_3 \mid r_1 \rangle \\ \langle r_1 \mid r_2 \rangle & \langle r_2 \mid r_2 \rangle & \langle r_3 \mid r_2 \rangle \\ \langle r_1 \mid r_3 \rangle & \langle r_2 \mid r_3 \rangle & \langle r_3 \mid r_3 \rangle \end{bmatrix} \tag{25}$$

and the 3×1 matrix $\beta(x)$,

$$\beta(x) = \begin{bmatrix} \langle r_1 \mid \widetilde{V}(x) \rangle \\ \langle r_2 \mid \widetilde{V}(x) \rangle \\ \langle r_3 \mid \widetilde{V}(x) \rangle \end{bmatrix}. \tag{26}$$

Then the projection is given by

$$\widehat{V}_a(x) = \Phi^{-1} \beta(x). \tag{27}$$

Using the matrix R previously defined, these equations can be expressed $\Phi = R^T W R$ and $\beta(x) = R^T W \widetilde{V}(x)$, so that

$$\widehat{V}_a(x) = (R^T W R)^{-1} R^T W \widetilde{V}(x)$$
$$= (R^T W R)^{-1} R^T W C_2 V(x). \tag{28}$$

Note that the 3×6 matrix $(R^T W R)^{-1} R^T W C_2$ is fixed and can be precomputed.

As mentioned previously, the resulting components $\widehat{V}_{aj}$ will not in general lie in the required interval [0,1]. One approach to perform normalization is as follows. Assume that the uniform anaglyph image $V_a(x) = V_{aw}$ gives the brightest achromatic (gray) image when viewed through the

anaglyph glasses. Ideally we would like $V_{aw} = [1\,1\,1]^T$. This is our best approximation of a uniform white $V_l(x) = V_r(x) = [1\,1\,1]^T$. Thus if we apply the projection equation (28) with $V(x) = E = [1\,1\,1\,1\,1\,1]^T$, we obtain $\widehat{E}_a = (R^T W R)^{-1} R^T W C_2 E$ instead of the desired V_{aw}. We can obtain the correct result by premultiplying by the diagonal normalizing matrix $N = \mathrm{diag}(V_{awj}/\widehat{E}_{aj})$. Thus with normalization included, the anaglyph image is given by

$$\widehat{V}_{an}(x) = N(R^T W R)^{-1} R^T W C_2 V(x)$$
$$= P V(x). \tag{29}$$

In this case the fixed 3×6 matrix $P = N(R^T W R)^{-1} R^T \cdot W C_2$ can be precomputed. The final step is clipping to the range [0,1] and application of gamma correction.

4. CONCLUSION

The projection method described in this paper has been used to generate numerous anaglyph images adapted to different types of colored glasses. The images produced are superior to those produced with classical methods such as the one described in [2]. The results cannot be shown in this paper since the images are meant to be viewed on a CRT display through colored glasses. Sample images can be found the the web page associated with this paper at the URL http://www.site.uottawa.ca/~edubois/icassp01/. Further work is required to find the optimal absorption curves for glasses matched to the display primaries and the properties of the human visual system, as well as to find a better error metric.

5. REFERENCES

[1] H. Gernshein and A. Gernshein, *The History of Photography from the Camera Obscura to the Beginning of the Modern Era.* New York, NY: McGraw-Hill, 1969.

[2] R. Turnnidge and D. Pizzanelli, "Methods of previsualizing temporal parallax suitable for making multiplex holograms. Part II: Greyscale and colour anaglyphs made in Photoshop™," *Imaging Science Journal*, vol. 45, no. 1, pp. 43–44, 1997.

[3] G. Wyszecki and W. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas.* New York, NY: John Wiley & Sons, 1967.

[4] C. Poynton, *A Technical Introduction to Digital Video.* New York, NY: John Wiley & Sons, 1996.

[5] A. Mertens, *Signal Analysis. Wavelets, Filter Banks, Time-Frequency Transforms and Applications.* Chichester, UK: John Wiley & Sons, 1999.