

RESEARCH

Open Access

Efficient 2D to 3D video conversion implemented on DSP

Eduardo Ramos-Diaz^{1*}, Victor Kravchenko² and Volodymyr Ponomaryov¹

Abstract

An efficient algorithm to generate three-dimensional (3D) video sequences is presented in this work. The algorithm is based on a disparity map computation and an anaglyph synthesis. The disparity map was first estimated by employing the wavelet atomic functions technique at several decomposition levels in processing a 2D video sequence. Then, we used an anaglyph synthesis to apply the disparity map in a 3D video sequence reconstruction. Compared with the other disparity map computation techniques such as optical flow, stereo matching, wavelets, etc., the proposed approach produces a better performance according to the commonly used metrics (structural similarity and quantity of bad pixels). The hardware implementation for the proposed algorithm and the other techniques are also presented to justify the possibility of real-time visualization for 3D color video sequences.

Keywords: disparity map, multi-wavelets, anaglyph, 3D video sequences, quality criteria, atomic function, DSP

1. Introduction

Conversion of available 2D content for release in three-dimensional (3D) is a hot topic for content providers and for success of 3D video in general. It naturally completely relies on virtual view synthesis of a second view given the original 2D video [1]. 3DTV channels, mobile phones, laptops, personal digital assistants and similar devices represent hardware, in which the 3D video content can be applied.

There are several techniques to visualize 3D objects, such as using polarized lens, active vision, and anaglyph. However, some of those techniques have certain drawbacks, mainly the special hardware requirements, such as the special display used with the synchronized lens in the case of active vision and the polarized display in the case of polarized lens. However, the anaglyph technique only requires a pair of spectacles constructed with red and blue filters where the red filter is placed over the left position producing a visual effect of 3D perception. Anaglyph synthesis is a simple process, in which the red channel of the second image (frame) replaces the red channel in the first image (frame) [2]. In the literature, several methods to compute anaglyphs have been

described. One of them is the original Photoshop algorithm [3], where the red channel of the left eye becomes the red channel of the anaglyph and vice versa for the blue and green channels of the right eye. Dubois [4] suggested the least square projection in each color component (R, G, B) from R_6 space to the 3D subspace. Two principal drawbacks of these algorithms are the presence of ghosting and the loss of color [5].

In the 2D to 3D conversion, depth cues are needed to generate a novel stereoscopic view for each frame of an input sequence. The simplest way to obtain 3D information is the use of motion vectors directly from compressed data. However, this technique can only recover the relative depth accurately, if the motion of all scene objects is directly proportional to their distance from the camera [1].

In [6], the motion vector maps, which are obtained from the MPEG4 compression standard, are used to construct the depth map of a stereo pair. The main idea here is to avoid the disparity map stage because it requires extremely computationally intensive operations and cannot suitably estimate the high-resolution depth maps in the video sequence applications. In paper [7], a real-time algorithm for use in 3DTV sets is developed, where the general method to perform the 2D to 3D conversion consists of the following stages: geometric analysis, static cues extraction, motion analysis, depth

* Correspondence: eramos@ieee.org

¹National Polytechnic Institute, ESIME-Culhuacan, Santa Ana 1000 Col. San Francisco Culhuacan, 04430, Mexico City, Mexico

Full list of author information is available at the end of the article

assignment, depth control, and depth image based rendering. One drawback of this algorithm is that it requires extremely computationally intensive operations.

There are several algorithms to estimate the DM such as the optical flow differential methods designed by Lucas & Kanade (L&K) and Horn and Schunk [8,9], where some restrictions in the motion map model are employed. Other techniques are based on the disparity estimation where the best match between pixels in a stereo pair or neighboring frames is found by employing a similarity measure, for example, the normalized cross-correlation (NCC) function or the sum of squared difference (SSD) between the matched images or frames [10]. A recent approach called the region-based stereo matching (RBSM) is presented in [11], where the block matching technique with various window sizes is computed. Another promising framework consists of stereo correspondence estimation based on wavelets and multi-wavelets [12], in which the wavelet transform modulus (WTM) is employed in the DM estimation. The WTM is calculated from the vertical and the horizontal detail components, and the approximation component is employed to normalize the estimation. Finally, the cross correlation in wavelet transform space is applied as the similarity measure.

In this article, we propose an efficient algorithm to perform a 3D video sequence from a 2D video sequence acquired by a moving camera. The framework uses the wavelet atomic functions (WAF) for the disparity map estimation. Then, the anaglyph synthesis is implemented in the visualization of the 3D color video sequence on a standard display. Additionally, we demonstrate the DSP

implementation for the proposed algorithm with different sizes of the 2D video sequences.

The main difference with other algorithms presented in literature is that the proposed framework performing sufficiently good depth and spatial perception in the 3D video sequences does not require intensive computational operations and can generate 3D videos practically in real-time mode.

In the present approach, we employ the WAFs because they have already demonstrated successful performance in medical image recognition, speech recognition, image processing, and other technologies [13-15].

The article is organized as follows: Section 2 presents the proposed framework, Section 3 contains the simulation results, and Section 4 concludes the article.

2. The proposed algorithm

The proposed framework consists of the following stages: 2D color video sequence decomposition, RGB component separation, DM computation using wavelets at multiple decomposition levels (M-W), in particular wavelet atomic functions (M-WAF), disparity map improvement via dynamic range compression, anaglyph synthesis employing the nearest neighbor interpolation (NNI), and 3D video sequence reconstruction and visualization. Below, we explain in detail the principal 3D reconstruction stages (Figure 1).

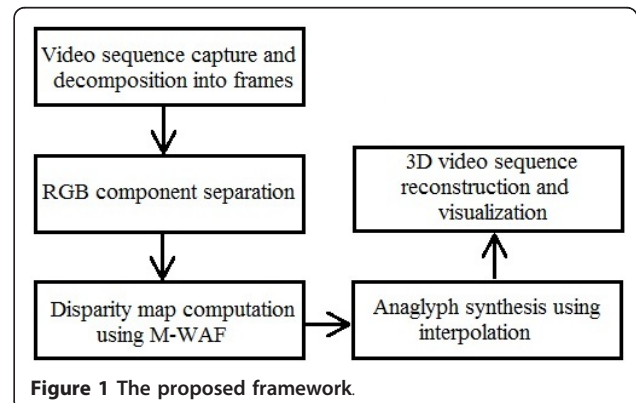
2.1. Disparity map computation

Stereo correspondence estimation based on the M-W (M-WAF) technique is proposed to obtain the disparity map. The stereo correspondence procedure consists of two stages: the WAF implementation and the WTM computation.

Here, we present a novel type of wavelets known as WAFs, first introducing basic atomic functions (up , fup_n , π_n) used as the mother functions in wavelet construction. The definition of AFs is connected with a mathematical problem: the isolation of a function that

Table 1 Filter coefficients $\{h_k\}$ for scale function $\phi(x)$ generated from different WAF based on up , fup_n , and π_n .

K	up	fup_n	π_n
0	0.757698251288	0.751690134933	0.7835967912
1	0.438708321041	0.441222946160	0.4233724330
2	-0.047099287129	-0.041796290935	-0.0666415128
3	-0.118027008279	-0.124987992607	-0.0793267472
4	0.037706980974	0.034309220121	0.0420426990
5	0.043603935723	0.053432685600	-0.0008988715
6	-0.025214528289	-0.024353106483	-0.0144489586
7	-0.011459893503	-0.022045882572	0.0211760726
8	0.013002207742	0.014555894480	-0.0046781803
9	-0.001878954975	0.007442614689	-0.0141324153
10	-0.003758906625	-0.006923189587	0.0104455879
11	0.005085949920	-0.001611566664	0.0003223058
12	-0.001349824585	0.002253528579	-0.0059986067
13	-0.003639380570	0.000052445920	0.0075295865
14	0.002763059895	-0.000189566204	-0.0011585840
15	0.001188712844	-0.000032923756	-0.0064315112
16	-0.001940226446	-0.000258206216	0.0047891344



has derivatives with a maximum and minimum similar to those of the initial function. To solve this problem requires an infinitely differentiable solution to the differential equations with a shifted argument [15]. It has been shown that AFs fall within an intermediate category between splines and classical polynomials: like B -splines, AFs are compactly supported, and like polynomials, they are universal in terms of their approximation properties.

The simplest and most important AF is generated by infinity-to-one convolutions of rectangular impulses that are easy to analyze via the Fourier transform. Based on N -to-one convolution of $(N + 1)$ identical rectangle impulses, the compactly supported spline $\theta_N(x)$ can be defined as follows:

$$\theta_N(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{jux} \left(\frac{\sin(u/2)}{u/2} \right)^{N+1} du. \quad (1)$$

The function $up(x)$ is represented by the Fourier transform for infinite convolutions of rectangular impulses with variable length of duration 2^{-k} , as in Equation 2:

$$up(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{jux} \prod_{k=1}^{\infty} \frac{\sin(u \cdot 2^{-k})}{u \cdot 2^{-k}} du. \quad (2)$$

The AF $fup_N(x)$ is defined by the convolution of spline $\theta_{N-1}(x)$ and AF $up(x)$ in the interval $[-(N+2)/2, (N+2)/2]$. Thus, $fup_N(x)$ can be written in the following form:

$$fup_N(x) = \int_{-\infty}^{\infty} e^{jux} \left(\frac{\sin(u/2)}{u/2} \right)^N \prod_{k=1}^{\infty} \frac{\sin(u \cdot 2^{-k})}{u \cdot 2^{-k}} du, \quad fup_0(x) \equiv up(x). \quad (3)$$

The generalization of AF $up(x)$ as presented above, the AF $up_m(x)$ is defined as follows:

$$up_m(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{jux} \prod_{k=1}^{\infty} \frac{\sin^2\left(\frac{mu}{(2m)^k}\right)}{\frac{mu}{(2m)^k} \sin\left(\frac{u}{(2m)^k}\right)} du, \quad m = 1, 2, 3, \quad up_1(x) = up(x). \quad (4)$$

The function $\pi_m(x)$ can be represented by the inverse Fourier transform $\pi_m(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ixt} F_m(t) dt$ using such representation for function $F_m(t)$:

$$F_m(t) = \prod_{k=1}^m \frac{\sin(2m-1)t + \sum_{v=2}^M (-1)^v \sin(2m-2v+1)t}{(3m-2)t}. \quad (5)$$

The detailed definitions and properties of these functions can be found in [15].

The wavelet decomposition procedures employ several decomposition levels to enhance the quality of the depth maps. The discrete wavelet transform (DWT) and inverse DWT are usually implemented using the filter

bank techniques for a scheme with only two filters: low pass (LP) $H(z)$ (decomposition) and $\tilde{H}(z)$ (reconstruction), and high pass (HP) $G(z)$ (decomposition) and $\tilde{G}(z)$ (reconstruction), where: $G(z) = zH(-z)$ and $\tilde{G}(z) = z^{-1}\tilde{H}(-z)$ [16]. The scale function $\phi(x)$ is associated with filter $H(z)$ in accordance to scaling equation:

$$\phi(x) = \frac{2}{H(1)} \sum_{k \in \mathbb{Z}} h_k \phi(2x - k) \text{ and can be expressed by}$$

$$\text{it Fourier transform } \hat{\phi}(\omega) = \prod_{k=1}^{\infty} \frac{H(e^{j\frac{\omega}{2^k}})}{H(1)}.$$

The wavelet functions are computed using linear combination of scale functions

$$\psi(x) = \frac{2}{H(1)} \sum_k g_k \phi(2x - k), \quad \text{where } g_k = (-1)^{k+1} h_{-k-1}^*,$$

and $\{h_k\}$ are the coefficients of the LP filter in it Fourier series:

$$H(\omega) = \sqrt{2}H_0(\omega) = \sum_k h_k e^{jk\omega} \quad \text{for } H_0(\omega) : h_k = \frac{\sqrt{2}}{2\pi} \int_{-\pi}^{\pi} H_0(\omega) e^{jk\omega} d\omega, \quad (6)$$

and wavelet $\tilde{\psi}(x) = \frac{2}{\tilde{H}(1)} \sum_k \tilde{g}_k \tilde{\phi}(2x - k)$. The HP filter is represented by Fourier series with coefficients $\{h_k\}$:

$$G(\omega) = e^{j\omega} H * (\omega + \pi) = \sum_k (-1)^{k+1} h_{-k-1}^* e^{-jk\omega}. \quad (7)$$

The coefficients $\{h_k\}$ should satisfy such normalization condition: $\frac{1}{\sqrt{2}} \sum_k h_k = H_0(0) = 1$. Finally, wavelets of decomposition and reconstruction are employed in such a form: $\tilde{\psi}_{i,k} = 2^{-i/2} \tilde{\psi}(x/2^i - k)$ and $\psi_{i,k} = 2^{-i/2} \psi(x/2^i - k)$, respectively, where i and k are indexes of translation and scale [16].

The procedure to synthesis the WAF consists of performing a scale function $\phi(x)$ that should generate the sequence of compact subspaces satisfying such property, each next subspace V_{j+1} is into a previous one V_j : $V_j \subset L^2(X)$, $j \in \mathbb{Z}$; $\cup_j V_j = L^2(X)$; $\cap_j V_j = \{0\}$; $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$. Finally, it should be existed such scale function $\phi(x)$ that: (a) with their shifts forms the *Riesz* bases; (b) it has symmetric and finite Fourier transform $\tilde{\phi}(\omega)$. Because the scale AF $\phi(x)$ and WAF $\psi(x)$ are not compactly supported but they rapidly decrease (due to infinite differentiability), it is possible to select an effective support from such limit conditions: $||\phi - \phi_{ef}|| \cdot 100\% \leq 0.001\%$, $||\psi - \psi_{ef}|| \cdot 100\% \leq 0.001\%$. Filter coefficients h_k for the scale function $\phi(x)$ generated from different WAFs: up , fup_n , up_n , π_n can be found in [17]. In Table 1, we only present the coefficients h_k for scale function $\phi(x)$ generated from AF up , fup_4 and π_6 that exposes better perception quality in synthesized 3D images as

one see below in simulation results. The effective supports for scale function $\phi(x)$ and wavelet $\psi(x)$ generated from used AF are $[-16, 16]$.

The Wavelet technique, which the developed method uses, is based on the DWT. In proposed framework for DM estimation, the wavelets on each decomposition level are computed as follows [12]:

$$W_s = |W_s| \angle \Theta_s, \quad (8)$$

$$|W_s| = \frac{\sqrt{|D_{h,s}|^2 + |D_{v,s}|^2 + |D_{d,s}|^2}}{|A_s|}, \quad (9)$$

where W_s is the wavelet for a chosen decomposition level s ; $D_{h,s}$, $D_{v,s}$, $D_{d,s}$ are the horizontal, vertical, and diagonal detail components at each a level s , A_s is the approximation component, and θ_s is the phase that is defined as follows:

$$\theta_s = \begin{cases} \varepsilon_s & \text{if } D_{h,s} > 0 \\ \pi - \varepsilon_s & \text{if } D_{h,s} < 0 \end{cases}, \quad \varepsilon_s = \arctg \frac{D_{h,s}}{D_{v,s}}. \quad (10)$$

Once the W_s is computed for each an image stereo pair or neighboring frames for a video, the disparity map for each level of decomposition can be formed using the cross-correlation function in wavelet transform space:

$$Cor_{(L,R),s}(x,y) = \sum_{(i,j) \in P} \frac{W_L(i,j) \cdot W_R(x+i,y+j)}{\sqrt{\sum_{i,j \in P} W_L^2(i,j) \cdot \sum_{i,j \in P} W_R^2(x+i,y+j)}}, \quad (11)$$

where W_L and W_R are the wavelet transform for the left and right images in each decomposition level s , and P is sliding processing window. Finally, the disparity map for each level of decomposition is computed by applying the NNI technique. In this work, we propose using four levels of decomposition in DWT.

A block diagram of the proposed M-WAF framework is presented in Figure 2.

2.2. Disparity map improvement and anaglyph synthesis

The classical methods used in anaglyph construction can produce ghosting effects and color loss. One way to reduce these artifacts in anaglyph synthesis is to use the dynamic range compression of the disparity map [18]. The dynamic range compression permits retaining the depth ordering information, which reduces the ghosting effects in the non-overlapping areas in the anaglyph. Therefore, the dynamic range reduction of the disparity map values can be employed to enhance the map quality. Using the P th law transformation for dynamic range compression [18], the original disparity map D is changed as follows:

$$D_{\text{new}} = a \cdot D^P, \quad (12)$$

where D_{new} is the new disparity map pixel value, $0 < a < 1$ is a normalizing constant, and $0 < P < 1$.

At the final stage, the anaglyph synthesis is performed using the improved disparity map. To generate an anaglyph, the neighboring frames in a grid dictated by the disparity map should be re-sampled. During numerous simulations, the *bilinear*, *sinc* and NNIs were implemented to find an anaglyph with a better 3D perception. The NNI showed a better performance during the simulations and it was sufficiently fast in comparison with the other investigated interpolations. Thus, the NNI was chosen to successfully create the required anaglyph in this application. The NNI is performed for each pair of neighboring frames in the video sequence. NNI [19] that uses this framework changes the values of the pixels to the closest neighbor value. To perform the NNI in the current decomposition level and to form the resulting disparity map, intensity of each pixel is changed. The new intensity value is determined by comparing a pixel in the low resolution disparity from i th decomposition level with the closest pixel value in the actual disparity map from $(i - 1)$ th decomposition level.

2.3. DSP implementation

Our study also involved employing the promising 3D visualization algorithms in real-time modes using a DSP. The core of the EVM DM642™ is a digital media processor that is characterized by a large set of integrated features of the card, such as: a TMS320DM642™ DSP at 720 MHz (1.39 instructions per cycle or 570 million instructions per second), 32 Mb of SDRAM, 4 Mb of Linear Memory Flash, 2 video decoders, 1 video coder, FPGA™ implementation to display, double UART with RS-232 drivers, several input-output video formats and others. The communication between the code composer studio (CCS) and the EVM is achieved with an external emulator via JTAG connectors [20]. Using MATLAB's Simulink™ module, a project was created in which the DSP model and its respective task BIOS were selected. Then, a function is created to contain three sub functions: video capture, 3D video reconstruction using WAF, and the output interface to a video display. Next, a CCS™ project is conducted in Simulink™. During this step in the process, the MATLAB™ module sends a signal to the CCS and creates the project on C. To perform the video sequence processing using the DSP, the MATLAB™ program is first transformed into 'C' code for CCS via Simulink™. Once the CCS project has been created, the necessary changes are made to obtain the processing time values. The corresponding results for the designed and the reference frameworks are presented in the next section. Serial connection of three EVM DM642 is used in this application, where the first and second DSPs compute the disparity maps using M-

Table 2 QBD and SSIM for proposed and existed algorithms for different test images

Image	L&K	SSD	GEEMSF	WF Bio6.8	WF Coiflet2	WF Haar	WAF π_6	M-WF Coiflet2	M-WAF π_6
Aloe									
SSIM	0.3983	0.6166	0.3017	0.9267	0.5826	0.5776	0.9232	0.5826	0.9232
QBP	0.1121	0.4722	0.9190	0.0297	0.4517	0.4420	0.0130	0.4490	0.0111
Venus									
SSIM	0.1990	0.4320	0.2145	0.5979	0.4530	0.4472	0.4604	0.4530	0.6947
QBP	0.3084	0.1428	0.2013	0.1694	0.5014	0.5010	0.1930	0.5011	0.1091
Lampshade1									
SSIM	0.0861	0.6320	0.3124	0.7061	0.7061	0.7081	0.6897	0.7061	0.7619
QBP	0.2430	0.2800	0.3410	0.2072	0.2071	0.2071	0.2017	0.2071	0.1426
Wood1									
SSIM	0.1089	0.7142	0.7051	0.9367	0.7096	0.7072	0.9448	0.7096	0.9448
QBP	0.1316	0.2376	0.2100	0.1258	0.2400	0.2402	0.1180	0.2400	0.0919
Bowling1									
SSIM	0.1118	0.6925	0.7081	0.8828	0.6690	0.6672	0.9084	0.6690	0.9084
QBP	0.1720	0.1885	0.0645	0.0555	0.2010	0.2011	0.0119	0.2010	0.0165
Reindeer									
SSIM	0.1557	0.7460	0.7143	0.7393	0.7321	0.7308	0.6819	0.7321	0.7001
QBP	0.3910	0.1250	0.2810	0.1418	0.1565	0.1570	0.1513	0.1520	0.1680

WAF procedure, and the third DSP generates the anaglyph. The developed algorithm in Simulink™ is shown in Figure 3.

3. Simulation results

In the simulation experiments, various synthetic images are used to obtain the quantitative measurements. The synthetic images were obtained from <http://vision.middlebury.edu/stereo/data>. *Aloe*, *Venus*, *Lampshade1*, *Wood1*, *Bowling1*, and *Reindeer* were the synthetic images used, all in PNG format (480 × 720 pixels). We also used the following test color video sequences in CIE format (250 frames, 288 × 352 pixels): *Coastguard*, *Flowers*, and *Foreman*. The test video sequences were obtained from <http://trace.eas.asu.edu/yuv/index.html>. In order to use the test color video sequences in the same sizes, we reformatted them in 480 × 720 pixels on Avi format. Additionally, the real life video sequences named *Video Test1* (200 frames, 480 × 720 pixels) and *Video Test2* (200 frames, 480 × 720 pixels) were recorded to apply the proposed algorithm in a common scenario. *Video Test1* shows a truck moving in the scenery and *Video Test2* shows three people walking toward the camera. Two quality objective criteria, quantity of bad disparities (QBD) [12] and similarity structure image measurement (SSIM) [21], were chosen as the quantitative metrics to justify the selection of the best disparity map algorithm in the 3D video sequence reconstruction. The QBD values have been calculated for different synthetic images as follows:

$$QBD = \frac{1}{N} \sum_{x,y} |d_E(x,y) - d_G(x,y)|^2, \quad (13)$$

where N is the total number of pixels in the input image, and d_E and d_G are the estimated and the ground truth disparities, respectively.

The SSIM metric values are defined as follows:

$$SSIM(x,y) = [l(x,y)] \cdot [c(x,y)] \cdot [s(x,y)], \quad (14)$$

where the parameters l , c , and s are calculated according to following equations:

$$l(x,y) = \frac{2\mu_X(x,y)\mu_Y(x,y) + C_1}{\mu_X^2(x,y) + \mu_Y^2(x,y) + C_1}, \quad (15)$$

$$c(x,y) = \frac{2\sigma_X(x,y)\sigma_Y(x,y) + C_2}{\sigma_X^2(x,y) + \sigma_Y^2(x,y) + C_2}, \quad (16)$$

$$s(x,y) = \frac{\sigma_{XY}(x,y) + C_3}{\sigma_X(x,y) + \sigma_Y(x,y) + C_3}. \quad (17)$$

In Equations (15) to (17), X is the estimated image, Y is the ground truth image, μ and σ are the mean value and standard deviation for the X or Y images, and $C_1 = C_2 = C_3 = 1$.

Table 2 presents the values of QBD and SSIM for the proposed framework based on M-WAFs and the other techniques applied to different synthetic images.

The simulation results presented in Table 2 indicate that the best overall performance of disparity map

Table 3 Processing times for different algorithms.

Algorithm	Matlab Time/frame, s (240 × 360)	Matlab Time/frame, s (480 × 720)	Serial Processing in DSP Time/frame, s (240 × 360)	Serial Processing in DSP Time/frame, s (480 × 720)
Classic wavelet families (<i>Coif2</i> , <i>Db6.8</i> , <i>Haar</i>)	4.20	6.16	0.0314	0.0713
Wavelet atomic functions (<i>up</i> , <i>fup_ϕ</i> , <i>π_ϕ</i>)	4.23	6.19	0.0312	0.0715
M-WAF (<i>up</i> , <i>fup_ϕ</i> , <i>π_ϕ</i>)	4.84	6.77	0.0489	0.081
M-classic wavelet families (<i>Coif2</i> , <i>Db6.8</i> , <i>Haar</i>)	4.85	6.76	0.0480	0.080

reconstruction is produced by the M-WAF framework. The minimum value of QBP and the maximum value of SSIM are obtained when the M-WAF π_6 is used, followed by WAF π_6 . At the final stage, when the anaglyphs were synthesized, the NCC was calculated in a sliding window with 5×5 pixels. The SSD algorithm was implemented in a window of size 9×9 pixels. The *L&K* algorithm was performed according to [9]. For all tested algorithms, the dynamic range compression was applied with the parameters $\alpha = P = 0.5$. Figure 4 shows the obtained disparity map for all tested images and all implemented algorithms; evidently, the M-WAF π_6 implementation produces the best overall visual results.

Based on the objective quantity metrics and the subjective results presented in Figure 4, M-WAF π_6 has

been selected as the technique to estimate the disparity map for video sequence visualization.

The anaglyphs, which were synthesized with the M-WAF algorithm, showed sufficiently good 3D visual perception with reduced ghosting and color loss. The spectacles with blue and red filters are required to observe Figures 5 and 6.

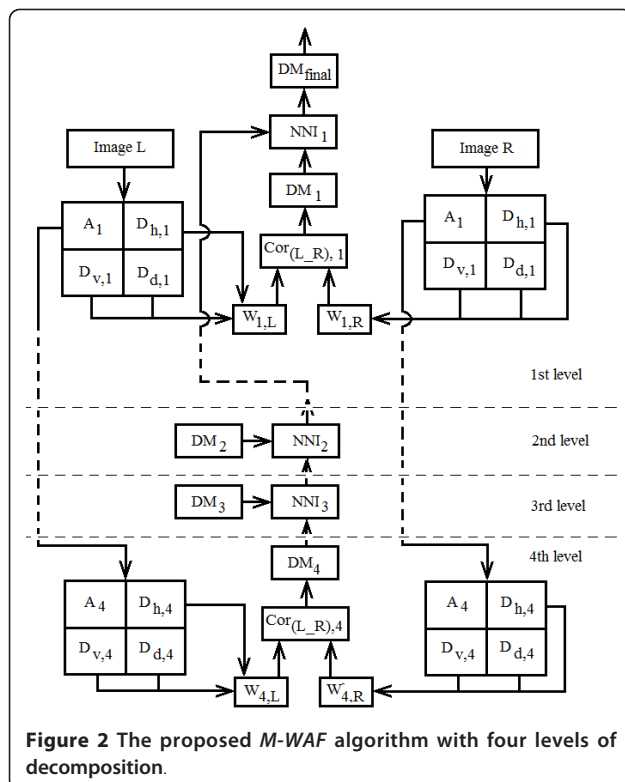
Processing time values were computed during the DSP implementation and the Table 3 shows the processing times for the video sequences using Matlab and the serial DSP implementation. Here, the tested video sequences were: *Flowers*, *Coastguard*, *Video Test1*, and *Video Test2* (all with 480×720 pixels and with 240×360 pixels in RGB format).

The processing time values were measured since the moment the sequence was acquired from the DSP until the anaglyph was displayed in a regular monitor.

The processing times in Table 3 lead to a possible conclusion that the DSP algorithm can process up to 20 frames per sec for a frame of 240×360 pixels size in RGB format. Additionally, the DSP algorithm can process up to 12 frames per sec for a frame of 480×720 pixels size in RGB format. Processing time values for *L&K* and SSD algorithms implemented in Matlab were 22.59 and 16.26 s, accordingly, because they required extremely computationally intensive operations.

4. Conclusion

This study analyzed the performance of various 3D reconstruction methods. The proposed framework based on M-WAFs is the most effective method to reconstruct the disparity map for 3D video sequences with different types of movements. Such framework produces the best depth and the best spatial perception in synthesized 3D video sequences against other analyzed algorithms that is confirmed by numerous simulations for different initial 2D color video sequences. The M-WAF algorithm can be applied to any type of color video sequence without additional information. The performance of the DSP implementation shows that the proposed algorithm can practically visualize the final 3D color video sequence in real-time mode. In future, we suppose to optimize the



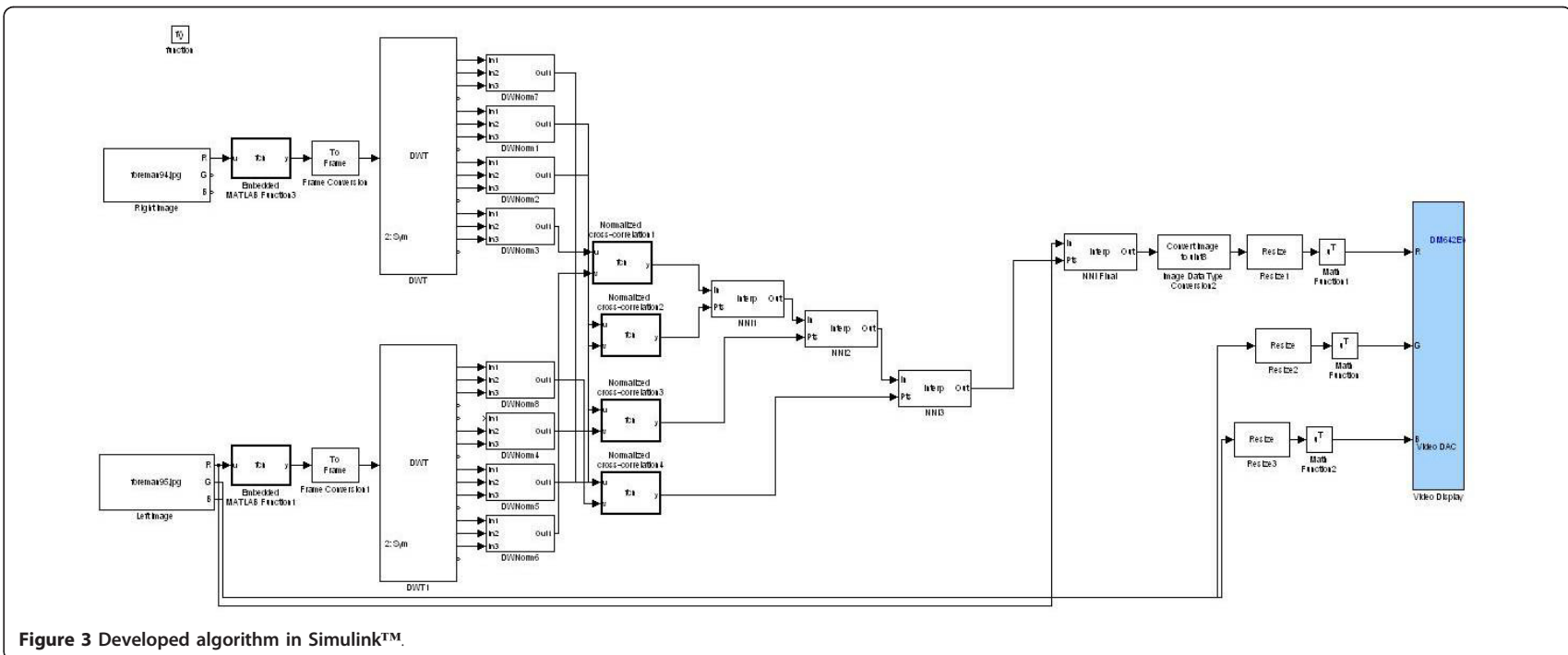


Figure 3 Developed algorithm in Simulink™.

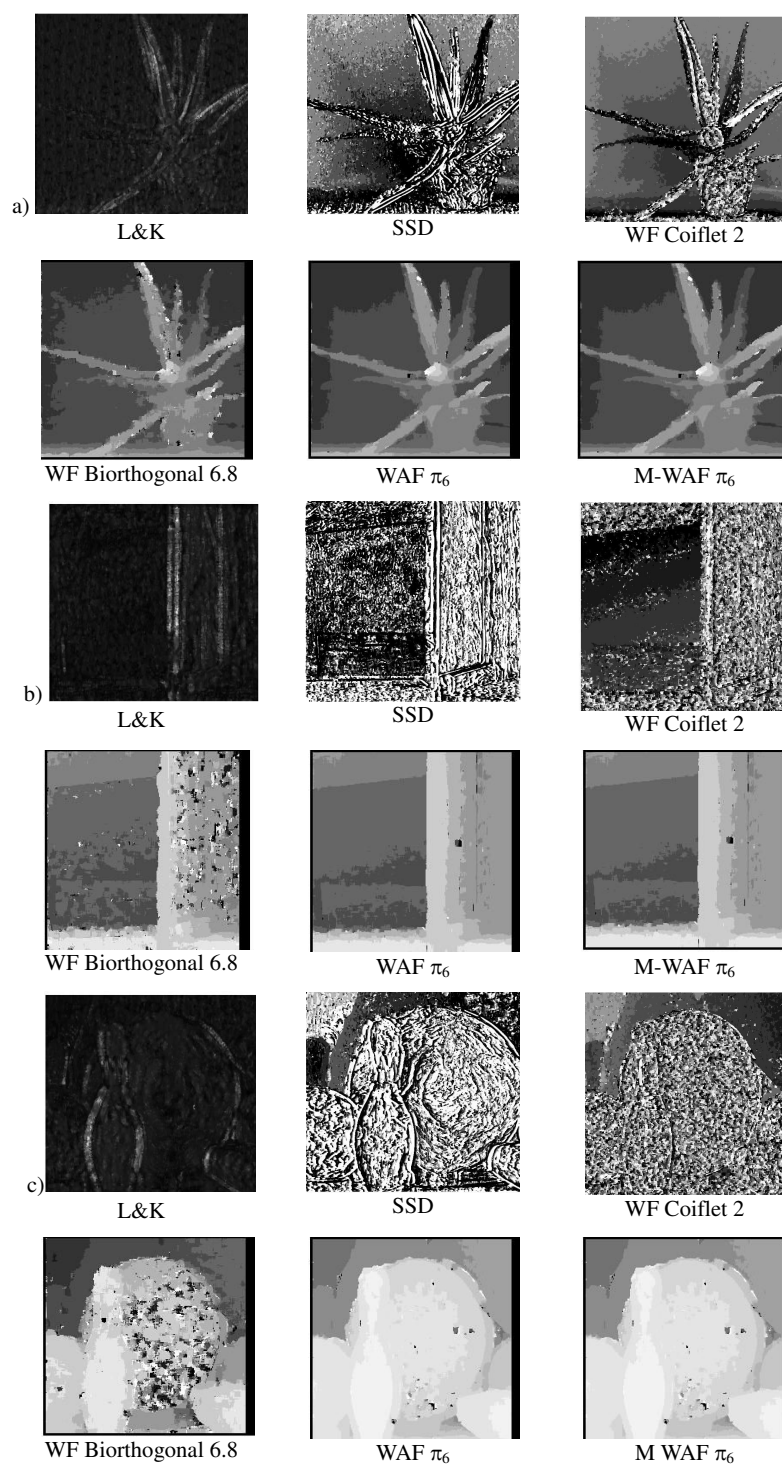


Figure 4 Disparity map obtained using different algorithms for following test images. (a) *Aloe*, (b) *Wood1*, and (c) *Bowling1*.



Figure 5 Synthesized anaglyphs using $M\text{-WAF } \pi_6$ for the following test images. (a) Venus, (b) Aloe, (c) Bowling1, (d) Lampshade, (e) Reindeer, and (f) Wood1.



Figure 6 Synthesized anaglyphs using $M\text{-WAF } \pi_6$ for frames of the following video sequences. (a) Flowers, (b) Coastguard, (c) Video Test1, and (d) Video Test2.

proposed algorithm in order to increase the processing speed up to the film velocity.

List of abbreviations

CCS: code composer studio; 3D: three-dimensional; LP: low pass; M-W: multiple decomposition levels; NCC: normalized cross-correlation; QBD: quantity of bad disparities; RBSM: region-based stereo matching; SSD: sum of squared difference; WAF: wavelet atomic functions; WTM: wavelet transform modulus.

Author details

¹National Polytechnic Institute, ESIME-Culhuacan, Santa Ana 1000 Col. San Francisco Culhuacan, 04430, Mexico City, Mexico ²Institute of Radio Engineering and Electronics, Russian Academy of Sciences, Moscow, Russia

Competing interests

The authors thank the National Polytechnic Institute of Mexico and CONACY (Project 81599) for their support of this work

Received: 3 June 2011 Accepted: 18 November 2011

Published: 18 November 2011

References

1. A Smolic, P Kauff, S Hnorr, A Hounnig, M Kunter, M Muller, M Lang, Three dimensional video postproduction and processing. *Proc IEEE*. **99**(4), 607–625 (2011)
2. I Ideses, L Yaroslavsky, New methods to produce high quality color anaglyphs for 3D visualization, in *ICIAR, Lecture Notes in Computer Science*, vol. 3212. (Springer Verlag, Germany, 2004), pp. 273–280. doi:10.1007/978-3-540-30126-4_34
3. W Sanders, D McAllister, Producing anaglyphs from synthetic images, in *Proceedings of SPIE Stereoscopic Displays and Virtual Reality Systems X*. **5006**, 348–358 (2003)
4. E Dubois, A projection method to generate anaglyph stereo images, in *Proceedings of IEEE International Conference on Acoustic Speech Signal Processing*, vol. 3. (Salt Lake City, USA, 2001), pp. 1661–1664
5. A Woods, T Rouke, Ghosting in anaglyphic stereoscopic images, in *Stereoscopic Displays and Applications XV, Proceedings of SPIE-IS&T Electronic Imaging, SPIE*. **5291**, 354–365 (2004)
6. I Ideses, L Yaroslavsky, B Fishbain, 3D from compressed video, in *Stereoscopic displays and virtual reality systems. Proc SPIE*. **6490**(64901C) (2007)
7. J Caviedes, J Villegas, Real time 2D to 3D conversion: Technical and visual quality requirements, in *International Conference on Consumer Electronics, ICCE-IEEE* 897–898 (2011)
8. DJ Fleet, *Measurement of Image Velocity*, (Kluwer Academic Publishers, Massachusetts, 1992)
9. SS Beauchemin, JL Barron, The computation of optical flow. *ACM Comput Surv.* **27**(3), 433–465 (1995). doi:10.1145/212094.212141
10. A Bovik, *Handbook of Image and Video Processing*, (Academic Press, USA, 2000)
11. BB Alagoz, Obtaining depth maps from color images by region based stereo matching algorithms. *OncuBilim Algor Syst Labs*. **08**(4), 1–12 (2008)
12. A Bhatti, S Nahavandi, in *Stereo Vision*, vol. Chap 6. (I-Tech, Vienna, 2008), pp. 27–48
13. Gulyaev YuV, VF Kravchenko, VI Pustovoi, A new class of WA-systems of Kravchenko-Rvachev functions in *Doklady mathematics*. **75**(2), 325–332 (2007)
14. C Juarez, V Ponomaryov, J Sanchez, V Kravchenko, Wavelets based on atomic function used in detection and classification of masses in mammography, in *Lecture Notes in Artificial Intelligence*. **5317**, 295–304 (2008)
15. V Kravchenko, H Meana, V Ponomaryov, in *Adaptive Digital Processing of Multidimensional Signals with Applications*, (FizMatLit Edit, Moscow, 2009) http://www.posgrados.esimecu.ipn.mx/
16. Y Meyer, *Ondelettes*, (Hermann, Paris, 1991)
17. VF Kravchenko, AV Yurin, New class of wavelet functions in digital processing of signals and images. *J Success Mod Radio Electron, Moscow, Edit Radiotekhnika*. **5**, 3–123 (2008)

18. I Ideses, L Yaroslavsky, Three methods that improve the visual quality of colour anaglyphs. *J Opt A Pure Appl Opt*. **7**, 755–762 (2005). doi:10.1088/1464-4258/7/12/008
19. A Goshtasby, *2D and 3D Image Registration*, (Wiley Publishers, USA, 2005)
20. Texas Instruments, TMS320DM642 Evaluation Module with TVP Video Encoders. Technical Reference 507345-0001 Rev B (December 2004)
21. WS Malpica, AC Bovik, Range image quality assessment by structural similarity. in *ICASSP 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE* 1149–1152 (2009)

doi:10.1186/1687-6180-2011-106

Cite this article as: Ramos-Diaz et al.: Efficient 2D to 3D video conversion implemented on DSP. *EURASIP Journal on Advances in Signal Processing* 2011 **2011**:106.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com