



HAL
open science

Novel View Synthesis for Stereoscopic Cinema: Detecting and Removing Artifacts

Frédéric Devernay, Adrian Ramos-Peon

► **To cite this version:**

Frédéric Devernay, Adrian Ramos-Peon. Novel View Synthesis for Stereoscopic Cinema: Detecting and Removing Artifacts. 3DVP 2010: ACM Workshop on 3D Video Processing, Oct 2010, Firenze, Italy. 10.1145/1877791.1877798 . inria-00531364

HAL Id: inria-00531364

<https://inria.hal.science/inria-00531364>

Submitted on 3 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Novel View Synthesis for Stereoscopic Cinema: Detecting and Removing Artifacts *

Frédéric Devernay
INRIA Grenoble
Rhône-Alpes, France
frederic.devernay@inrialpes.fr

Adrian Ramos Peon
INRIA Grenoble
Rhône-Alpes, France
adrian.ramos-peon@inrialpes.fr

ABSTRACT

Novel view synthesis methods consist in using several images or video sequences of the same scene, and creating new images of this scene, as if they were taken by a camera placed at a different viewpoint. They can be used in stereoscopic cinema to change the camera parameters (baseline, vergence, focal length...) *a posteriori*, or to adapt a stereoscopic broadcast that was shot for given viewing conditions (such as a movie theater) to a different screen size and distance (such as a 3DTV in a living room) [3]. View synthesis from stereoscopic movies usually proceeds in two phases [11]: First, disparity maps and other viewpoint-independent data (such as scene layers and matting information) are extracted from the original sequences, and second, this data and the original images are used to synthesize the new sequence, given geometric information about the synthesized viewpoints. Unfortunately, since no known stereo method gives perfect results in all situations, the results of the first phase will most probably contain errors, which will result in 2D or 3D artifacts in the synthesized stereoscopic movie. We propose to add a third phase where these artifacts are detected and removed in each stereoscopic image pair, while keeping the perceived quality of the stereoscopic movie close to the original.

Categories and Subject Descriptors

I.4.8 [Image Processing and computer vision]: Scene Analysis—*Stereo*

General Terms

Algorithms

Keywords

Stereoscopic Cinema, View Interpolation, Anisotropic filtering

*Work done within the 3DLive project supported by the French Ministry of Industry <http://3dlive-project.com/>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

3DVP'10, October 29, 2010, Firenze, Italy.

Copyright 2010 ACM 978-1-4503-0159-6/10/10 ...\$10.00.

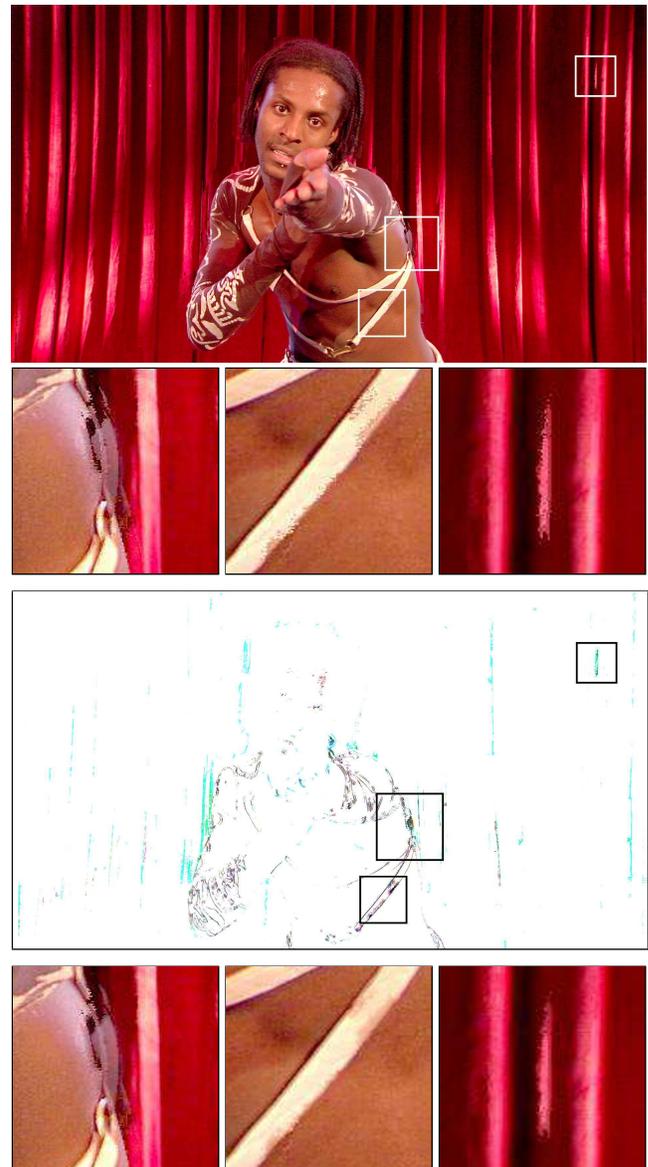


Figure 1: Top row: synthesized novel view with zoom on three artifacts. Middle row: confidence map used to detect artifacts. Bottom row: results of artifact removal.

1. INTRODUCTION

Novel view synthesis consists in taking several images or videos of the same scene with a number of synchronized cameras, and creating a new image or video as if it were taken from a new viewpoint which is not present in the original set of camera positions. It has been the subject of lots of recent research, sometimes with quite impressive results [11, 1, 4, 5, 10].

In the context of stereoscopic cinema or 3DTV, the set of cameras is usually reduced to two, and novel view synthesis is used to generate a new stereoscopic pair of videos, in order to change the perceived scene geometry, or to adapt to the display size. Most of these movies are shot and produced for a fixed display size and viewing distance (usually a typical movie theater screen). However, to be able to display the same movie on different displays (e.g. in a non-standard movie theater, on a home cinema system, or on a smaller 3DTV), simple scaling of the images is not enough, since it distorts the image in the depth direction, due to the fact that the human interocular is a fixed constant in the display geometry and that the perceived depth is also proportional to the distance to the screen [2, 3].

Novel view synthesis for stereoscopic movies is thus a sub-problem of the multi-view case, with less data available and more constraints on the synthesized viewpoints [7, 3]. Rogmans *et al.* [7] reviewed the available methods for novel view synthesis from stereoscopic data, and noticed that they essentially consist of two steps: first, a stereo correspondence module computes the stereoscopic disparity between the two views, and second, a view synthesis module generates the new views, given the results of the first module and the parameters of the synthesized cameras. The main consequence is that any error in the first module will generate artifacts in the generated views. These can either be 2D artifacts, which appear only on one view and may disrupt the perceived scene quality and understanding, or even worse: 3D artifacts, that may appear as floating bits in 3D and look very unnatural.

We thus propose to add a third module that will detect artifacts, and remove them by smoothing them out. The key idea is that stereoscopic novel view synthesis can be done in an asymmetric way. As noted by Seuntings *et al.* [8], if one of the views is close to or equal to the original image, the other view can be slightly degraded without any negative impact on the perceived quality of the stereoscopic movie, and eye dominance has no effect on the quality. We thus propose to use *asymmetric novel view synthesis*, where the left view is the original image, and only the right view is synthesized. Consequently, artifacts are only present in the right view, and we propose to detect and remove them by smoothing. In these very small modified areas of the stereoscopic image pair, the visual system will use the left view combined with 3D cues other than stereopsis to reconstruct the proper 3D geometry of the scene.

The remaining of this paper is organized as follows: Sec. 2 describes the basic algorithm used to synthesize the novel right view in the simple case of *baseline modification* (however, more general novel view synthesis methods can be used, as long as they preserve the epipolar geometry of the stereo pair [3]). Sec. 3 describes the artifact detection and removal algorithm, which is based on building a confidence map on the synthesized view, and then applying anisotropic diffusion on that view based on the Perona-Malik equation. Fi-

nally, Sec. 4 presents results and concludes on the necessary psycho-visual evaluation of this method and possible extensions.

2. NOVEL VIEW SYNTHESIS

In the context of novel view synthesis from a stereoscopic pair of images or videos, one usually assumes that the original images were rectified, so that there is no vertical disparity between the two images: epipolar lines are horizontal, and a point (x_l, y) in the left image corresponds to a point at the same y coordinate (x_r, y) in the right image. The 3D information about the part of the scene that is visible in the stereo pair is fully described by the camera parameters, and the disparity maps that describe the mapping between points in the two images.

Let $I_l(x_l, y)$, $I_r(x_r, y)$ be a pair of rectified images, and $d_l(x_l, y)$, $d_r(x_r, y)$ be respectively the left-to-right and right-to-left disparity maps: d_l maps a point (x_l, y) in the left image to the point $(x_l - d_l(x_l, y), y)$ in the right image, and d_r maps a point (x_r, y) in the right image to the point $(x_r + d_r(x_r, y), y)$ in the left image (signs are set so that the bigger the disparity, the closer the point). These two disparity maps may be produced by any method, and the semi-occluded areas, which have no correspondent in the other image, are supposed to be filled using some assumption on the 3D scene geometry. A stereoscopic pair of images and their corresponding disparity maps used in our examples are shown in Fig. 2. These were produced by the crudest stereo method: block matching with a fixed square window size, winner-takes-all disparity selection, left-right cross-checking, hole filling using the highest depth found around each hole, and basic smoothing. Of course, this method produces large errors which will result in visible interpolation artifacts and help demonstrate the usefulness of our method, but this could be replaced by any state-of-the-art stereo matching method. A better stereo method will produce less visible 2D artifacts in the interpolated views, but when viewed on a stereoscopic display the scene may still contain 3D artifacts. The sources of matching errors which will cause artifacts are mainly non-textured or out-of-focus areas, depth discontinuities, repetitive patterns, and specular reflections.

2.1 Backward mappings

In the synthesized view, each pixel may have a visible matching point in the left image and/or a visible correspondent in the right image. If the point is not visible in one of the original images, the mapping is undefined at that point. We call these mappings from the synthesized view to the original images *backward mappings*. If the desired output is a stereoscopic pair, there may be two synthesized views and two sets of backward mappings to the original images. However, as explained in the introduction, we focus on asymmetric synthesis methods, where the left image in the output stereoscopic pair is the original left image, and only the right image in the output stereoscopic pair is synthesized. Different methods are available for asymmetric synthesis, producing various effects on the perceived 3D geometries [3]. The simplest one is *asymmetric baseline modification*, where the synthesized right image corresponds to a viewpoint placed on the baseline, which is the straight line joining the original viewpoints, between these two viewpoints. In the following, we derive equations for this interpolation method, but any

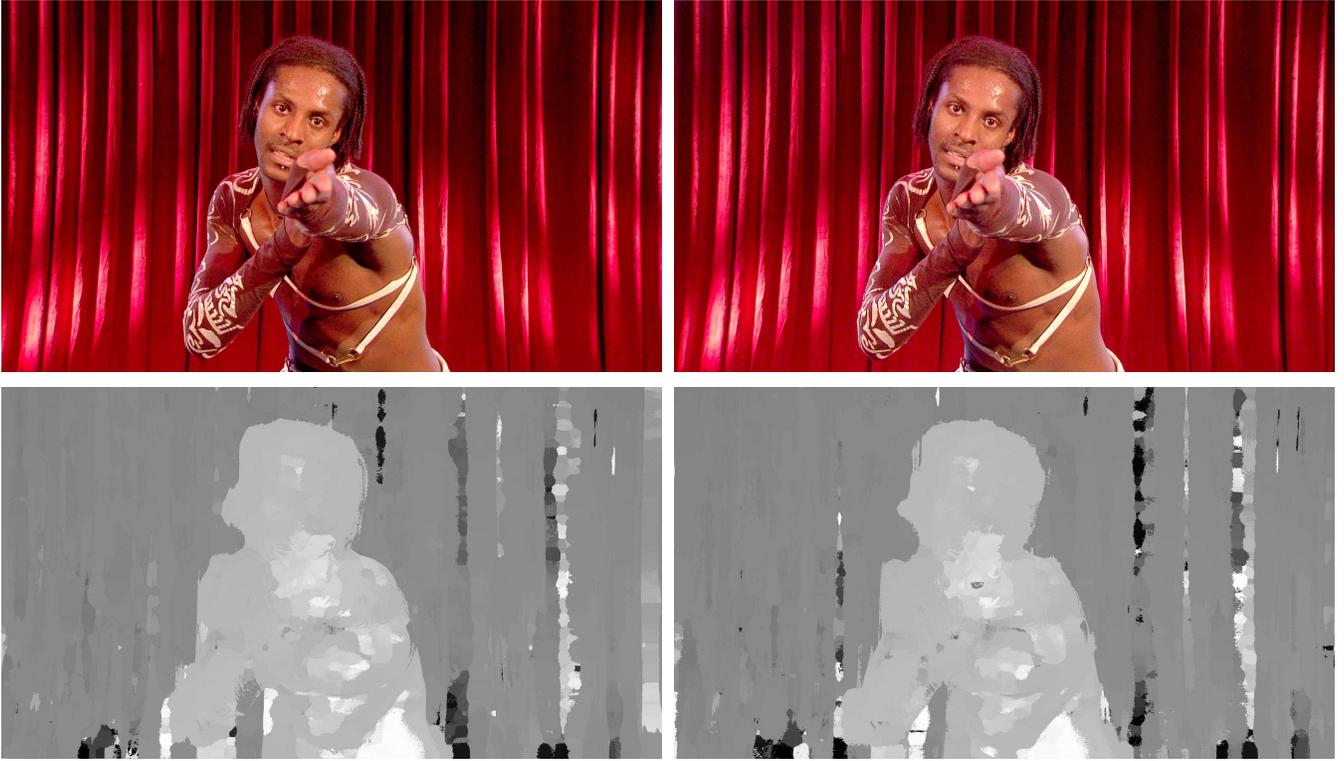


Figure 2: Top row: the left and right rectified images from the original stereo pair (Liège - 2009, Images courtesy of RTBF / Outside Broadcast / Binocle). Bottom row: the left and right disparity maps produced from these images by a basic method, obviously containing many matching errors.

method will work, provided that the backward mappings to the original images are available or computable.

Let $I_\alpha(x_\alpha, y)$ be the synthesized image corresponding to the interpolation position α ($0 \leq \alpha \leq 1$), between the left viewpoint ($\alpha = 0$) and the right one ($\alpha = 1$). Since the left image is unchanged and the output stereo pair must be rectified, the backward mapping from the synthesized viewpoint to each original image only has a horizontal component: it is a backward disparity map.

Let d_α^l and d_α^r be the backward disparity maps from the interpolated viewpoint, respectively to the left and to the right original images (the subscript is the reference viewpoint, and the superscript is the destination image). d_α^l and d_α^r map each integer-coordinates point in I_α to a real-coordinates point in I_l and I_r , or to an undefined value if the point is not visible in the corresponding original image.

Let us describe how backward maps are computed from the original disparity maps. Multiplying respectively d_l and d_r by the scalars α and $1 - \alpha$ produces *forward* maps ϕ_l^α and ϕ_r^α from the original images to the interpolated viewpoint (a map from viewpoint a to viewpoint b is noted ϕ_a^b). Since all these mappings only act on the x coordinate, we can forget the y coordinate in the following:

$$I_l \rightarrow I_\alpha : x_l \mapsto x_l - \alpha d_l(x_l, y) = \phi_l^\alpha(x_l, y) \quad (1)$$

$$I_r \rightarrow I_\alpha : x_r \mapsto x_r + (1 - \alpha)d_r(x_r, y) = \phi_r^\alpha(x_r, y) \quad (2)$$

Backward mapping is computed by inverting the piecewise linear extensions of the discrete functions ϕ_l^α and ϕ_r^α . First, the discrete array of the backward map $\phi_\alpha^l(x_\alpha)$ is initialized to an invalid value for all x_α . Then, we consider that a

linear segment exists between discrete parameter values x_l and $x_l + 1$ of ϕ_l^α iff $0 < \phi_l^\alpha(x_l + 1) - \phi_l^\alpha(x_l) < e$, where e is an expansion threshold, typically set to 2 pixels, used to distinguish depth discontinuities from slanted surfaces. $\phi_\alpha^l(x_\alpha)$ is then computed by inverting the resulting function:

$$\forall x_l, \forall x_\alpha \in [[\phi_l^\alpha(x_l)], [\phi_l^\alpha(x_l + 1)]], \phi_\alpha^l(x_\alpha) = \frac{x_\alpha - b}{a}, \quad (3)$$

where $a = \phi_l^\alpha(x_l + 1) - \phi_l^\alpha(x_l)$ and $b = \phi_l^\alpha(x_l + 1) - a(x_l + 1)$. $\phi_\alpha^r(x_\alpha)$ is computed similarly.

The simplest way to handle occlusions is by using a z -buffer algorithm: each time a new value for $\phi_\alpha^l(x_\alpha)$ or $\phi_\alpha^r(x_\alpha)$ is computed, if a valid value already exists, it is overwritten only if the new value corresponds to a closer point (i.e. a larger disparity). However, we notice that the z -test can be avoided by using the painter's algorithm: if d_l is swept from left to right to build ϕ_α^l , and d_r is swept from right to left to build ϕ_α^r , newer values are always closer and overwrite the previously computed ones.

2.2 Image synthesis

With the backward disparity maps, we can now get any kind of value for (almost) all pixels in the synthesized viewpoint, be it intensity, Laplacian or gradient, as long as it can be computed in the left and right images. To get $I_\alpha(x_\alpha, y)$, the pixel intensity at (x_α, y) , we will begin by finding $I_\alpha^l(x_\alpha, y)$ and $I_\alpha^r(x_\alpha, y)$, that is, the intensities in the left and right images corresponding to each point (x_α, y) in the novel view. These are computed by linear interpolation of the intensities

at position $\phi_\alpha^l(x_\alpha, y)$ of the values in I_l , and at $\phi_\alpha^r(x_\alpha, y)$ in I_r .

However, the values of $\phi_\alpha^l(x_\alpha, y)$ and $\phi_\alpha^r(x_\alpha, y)$ might be invalid due to disocclusion. A pixel in the synthesized view may not be visible in one or both of the original viewpoints, although no hole is present in the initial disparity maps. Several strategies exist for hole handling (see Rogmans *et al.* [7] for a review), but we opted for a simple method. We have three possible cases: only one of the values is valid, both are valid, or none is valid. In the first case, the resulting intensity of $I_\alpha(x_\alpha, y)$ will be that fetched by the valid mapping. If both $I_\alpha^l(x_\alpha, y)$ and $I_\alpha^r(x_\alpha, y)$ can be retrieved, then they are blended according to α , and yield:

$$I_\alpha(x_\alpha) = (1 - \alpha)I_\alpha^l(x_\alpha) + \alpha I_\alpha^r(x_\alpha). \quad (4)$$

In more complex novel view synthesis methods, the blending factor may also depend on the disparity itself [3].

If the backward disparity maps in both directions have a hole at the same pixel, no value can be retrieved. This would be the case when the interpolated view shows a region that isn't available in any of the two original viewpoints. It is a serious problem, since in our setting there is no way to retrieve this information, which could only be available by using an additional camera. Some elaborate inpainting solutions exist to fill these zones according to different criteria [9]. However this is outside the scope of our problem, and therefore, instead of leaving blank holes, we just compute the value at these pixels by taking the nearest available neighbors on the same horizontal line and interpolate linearly. The synthesized intermediary view in between those shown in Figure 2 can be seen in Figure 4. Artifacts can be seen at various places. Some of them are highlighted in Figure 1.

3. ARTIFACT DETECTION AND REMOVAL

The quality of the synthesized view highly depends on the quality of the disparity maps. For example, the disparity maps shown in Fig. 2, together with the novel view synthesis method described in the previous section, generate the interpolated image shown on top row of Fig. 1, which contains numerous small artifacts. These artifacts are usually small high-frequency defects, and when viewed on a stereoscopic display, they usually appear as floating bits or holes in the surface, and even an inexperienced viewer easily notices them. Besides, when novel view synthesis is done on a stereoscopic movie, these artifacts may appear and disappear from frame to frame, resulting in a very disturbing flickering effect.

Since the stereo correspondence module usually implements the best possible algorithm given the system constraints¹, the disparity maps cannot be improved. Instead, we propose a method to detect the areas in the images where these artifacts are present, and then remove them by smoothing. That way, since the left image in our asymmetric view synthesis method is the high-quality original image, only this image will be used by the visual system to perceive the 3D scene geometry in the small areas covered by these artifacts.

¹To produce our results, however, we used a low-quality stereo algorithm to demonstrate the capabilities of our method.

3.1 Artifact detection

Artifact detection works by building a confidence map over the whole interpolated image, where most pixels are marked with high confidence, and artifacts are marked with low confidence.

Once an interpolated image has been generated, we want to create a confidence map, attributing a weight to each pixel to specify how certain we are about its correctness.

Having the original left and right viewpoints of a scene, and the interpolated viewpoint of the same scene, building this confidence map is based on the fact that we expect to find similar pixel intensities, gradients and Laplacians in all images (excluding occlusions), but at different locations due to the geometric mappings between these views. Using this observation, we are able to outline areas and edges which should not appear in the interpolated view. For instance, a gradient appearing in the synthesized view that does not exist in either of the two original views should suggest the presence of an artifact, and will be marked as a low confidence zone in our confidence map.

In fact, the backward mappings ϕ_α^l and ϕ_α^r defined in the previous section can also be used to compare the intensities, gradients, and Laplacians of the interpolated view with the left and right views. Theoretically, warped derivatives (gradients and Laplacian) should be composed with the derivatives of the mappings, but we make the assumption that the scene surface is locally fronto-parallel and ignore these, because the mappings derivatives contain too much noise. At first, three separate confidence maps can be created, based on these different values (intensity, gradient, Laplacian). They are computed in the following way:

1. Get the value of the pixel on the interpolated image.
2. Fetch the values of the corresponding pixels (according to backward mappings) in the left and right images (if available)
3. Compute the absolute value of the difference with each value (if available).
4. If values from the left and right images are available, then blend the absolute differences according to α , else take the only available value.

The artifacts that appear in the synthesized view are mainly composed of high frequency components of the image, thus the Laplacian differences should give a good hint on where the artifacts are located. Intensity or gradient differences, on the other hand, may appear at many different places which are not actual artifacts, such as large specular reflections, or intensity differences due to the difference in illumination.

Using the Laplacian only works great on images with high texture and fine details, such as those found in the Middlebury stereo benchmark dataset. These only cause small artifacts, such as lines or spots, which can easily be detected by the Laplacian. However, with HD or higher resolution images, the area of each artifact (in pixels) becomes larger, and the Laplacian only detects its contour, and not the inner region. These regions are better detected with the intensity differences, although not perfectly, since erroneous disparity maps may point to other pixels with the same color. Also, since the intensity in both original images can vary significantly due to differences in the camera settings and parameters, as well as specular reflection zones, the confidence map

built from intensities is subject to big variations and is not very robust.

We thus want to detect artifacts as areas which are surrounded by high Laplacian differences and inside which the intensity or gradient difference with the original images is high. We start by dilating the Laplacian using a small structuring element (typically a 3×3 square), so that not only the borders are detected, but also more of the inner (and outer) region of the artifact. Then, to remove the regions outside the actual artifacts (introduced by the dilation), we multiply this dilated map by the intensity difference map. This multiplication partly alleviates the specularity problem, since the regions detected as “uncertain” by the intensity map alone are now compared using the Laplacian, and if there are no discrepancies in the Laplacian differences, this area will be marked as being correct in the resulting confidence map. To further avoid incorrect detections introduced by the intensity differences, we decide to discard the weakest values. To do so, we set a threshold so that at most 5% of the image will have a non-zero value in the confidence map. This prevents overall image blurring in subsequent treatment of the interpolated image. The confidence map obtained in this way on the sample stereo pair and synthesized view is shown in Figure 1 (middle row) and details are shown in Figure 3. As can be seen, larger artifacts are indeed well detected.

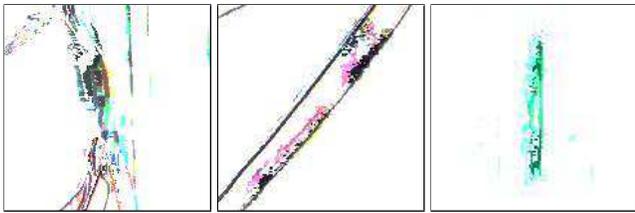


Figure 3: Zoom of the color inverted confidence map on the artifacts from Fig. 1 (white corresponds to zero in the confidence map, representing correct values).

3.2 Artifact removal by anisotropic blurring

To try to remove as much as possible of the detected artifacts, we use the anisotropic diffusion equation, as used by Perona-Malik [6]:

$$\frac{\partial I}{\partial t} = \nabla \cdot (c(x, y, t) \nabla I) = c(x, y, t) \Delta I + \nabla c \cdot \nabla I \quad (5)$$

where $c(x, y, t)$ are the conduction coefficients, which guide the smoothing of the image. While Perona and Malik were interested in finding the coefficients to smooth the image only within regions, and not across boundaries, we already know where we want the diffusion to occur. The confidence map provides these space variant coefficients, that cause the detected artifacts to be smoothed out, while the rest of the image is left untouched. The values from the confidence map are normalized so that the resulting coefficients are between zero and one.

The numerical scheme used to implement the proposed smoothing is simple. At each iteration, the intensity at each pixel is changed according to the following rule, producing

a new image which is used for the subsequent iteration:

$$I^{t+1}(x, y) = \frac{\Delta t}{2} \left((c(x-1, y) + c(x, y))(I^t(x-1, y) - I^t(x, y)) \right. \\ \left. + (c(x+1, y) + c(x, y))(I^t(x+1, y) - I^t(x, y)) \right. \\ \left. + (c(x, y-1) + c(x, y))(I^t(x, y-1) - I^t(x, y)) \right. \\ \left. + (c(x, y+1) + c(x, y))(I^t(x, y+1) - I^t(x, y)) \right). \quad (6)$$

Notice that the t is dropped in the conduction coefficients, since they are constant in time. Having the confidence map drive the diffusion has the effect that at each iteration, the intensities surrounding the artifact propagate inwards, while the contrary does not hold. After a sufficient number of iterations, the artifact reduces in size.

Following the reasoning found in [6], it is seen that more complex implementations of the diffusion equation increase the computational complexity, but however yield perceptually similar results. Since most of the coefficients from the confidence map are zero, the cost of the diffusion is very small compared to a full image blurring. Moreover, the diffusion coefficients being constant in time, they do not have to be re-computed at each iteration, sparing additional operations from the conventional Perona-Malik algorithm.

4. RESULTS

We show results on one frame of a stereoscopic movie². The input data to our novel view synthesis and artifact removal modules are the left and right images, as well as the left and right *dense* disparity maps (Fig. 2). The interpolated viewpoint is synthesized at $\alpha = 0.5$, i.e. between the left and right viewpoints, where more artifacts are present (since it is farther apart from any original viewpoint). The confidence maps are created by performing one dilation of the Laplacian confidence map using a 3×3 structuring element before multiplying it with the intensity confidence map. The anisotropic blurring is performed with a time step of $\Delta t = 0.3$, and 20 iterations are computed.



Figure 4: The synthesized right image, after artifact removal (compare with top row of Fig. 1).

The right synthesized view is shown on Fig. 4, and zoom on details is available on bottom row of Fig. 1. Small and

²The full stereoscopic sequence - original, synthesized, and with artifacts removed - will be made available before the workshop if a standard format for stereoscopic movies is chosen by the workshop organization

medium artifacts were detected and removed by our algorithm, but some of the bigger artifacts are still present. We notice for example on Fig. 1 that the “curtain” artifact was not completely removed because there is a very large matching error in the original disparity maps, due to a repetitive pattern with slight occlusions (the curtain folds), and part of the resulting artifact is consistent with the original images and disparity maps, as can be seen in the confidence map (Fig. 3). It proves that the disparity maps still have to be of an acceptable quality in order to remove properly all the artifacts, and the final quality of the stereo pair still depends on the quality of the stereo correspondence module, although in lesser proportions than if this artifact removal module is not present: a state-of-the-art stereo correspondence methods will produce less and smaller artifacts which will be easily removed by the proposed method (but the artifacts would be almost unnoticeable on a monoscopic image, although they still appear when viewed in 3D).

Some of the natural artifacts that should be present in the synthesized image, such as specular reflections in the eyes, were also smoothed a little, but the impact on the resulting perceived quality of the stereoscopic pair is not important, since the left image still has these natural artifacts (specular reflections do not follow the epipolar constraint, and are thus rarely matched between the two views, even in the human visual system, although they still bring a curvature cue on the local surface geometry).

5. CONCLUSION AND FUTURE WORK

Novel view synthesis methods for stereoscopic video usually rely on two algorithmic modules which are applied in sequence to each stereoscopic pair in the movie [7]: a stereo correspondence module and a view synthesis module. Unfortunately, in difficult situations such as occlusions, repetitive patterns, specular reflections, low texture, optical blur, or motion blur, the stereoscopic correspondence module produces errors which appear as artifacts in the final synthesized stereo pair. We propose to add a third module, which detects artifacts in the synthesized view by producing a confidence map, and then smooths out this artifacts by anisotropic diffusion based on the Perona-Malik equation [6].

The results show that this method removes small artifacts from the synthesized view. However, large artifact that are consistent both with the original images and the disparity maps may remain after this process, so the quality of the stereo correspondence module is still crucial for artifact-free novel view synthesis.

Since these preliminary results are promising, we intend to work on the validation of this method by a psycho-visual study involving several viewers, in order to evaluate quantitatively the quality improvements brought by the application of this artifact removal module. We also plan to work on integrating temporal consistency by the combined use of disparity maps and optical flow maps, in order to reduce the appearance of flickering artifacts in stereoscopic movies, which are probably the most disturbing spatial artifacts for the movie viewer.

6. REFERENCES

- [1] A. Criminisi, A. Blake, C. Rother, J. Shotton, and P. H. Torr. Efficient dense stereo with occlusions for new view-synthesis by four-state dynamic programming. *Int. J. Comput. Vision*, 71(1):89–110, 2007.
- [2] F. Devernay and P. Beardsley. Stereoscopic cinema. In R. Ronfard and G. Taubin, editors, *Image and Geometry Processing for 3-D Cinematography*. Springer-Verlag, 2010.
- [3] F. Devernay and S. Duchêne. New view synthesis for stereo cinema by hybrid disparity remapping. In *International Conference on Image Processing (ICIP)*, Hong Kong, 2010.
- [4] D. Farin, Y. Morvan, and P. H. N. de With. View interpolation along a chain of weakly calibrated cameras. In *IEEE Workshop on Content Generation and Coding for 3D-Television*, 2006.
- [5] J. Kilner, J. Starck, and A. Hilton. A comparative study of free-viewpoint video techniques for sports events. In *Proc. 3rd European Conference on Visual Media Production*, pages 87–96, London, UK, 2006.
- [6] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.
- [7] S. Rogmans, J. Lu, P. Bekaert, and G. Lefruit. Real-time stereo-based view synthesis algorithms: A unified framework and evaluation on commodity GPUs. *Signal Processing: Image Communication*, 24(1-2):49–64, 2009. Special issue on advances in three-dimensional television and video.
- [8] P. Seuntjens, L. Meesters, and W. Ijsselstein. Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation. *ACM Trans. Appl. Percept.*, 3(2):95–109, Apr. 2009.
- [9] L. Wang, H. Jin, R. Yang, and M. Gong. Stereoscopic inpainting: Joint color and depth completion from stereo images. *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8, June 2008.
- [10] O. Woodford, I. D. Reid, P. H. S. Torr, and A. W. Fitzgibbon. On new view synthesis using multiview stereo. In *Proceedings of the 18th British Machine Vision Conference*, volume 2, pages 1120–1129, Warwick, 2007.
- [11] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski. High-quality video view interpolation using a layered representation. In *Proc. ACM SIGGRAPH*, volume 23, pages 600–608, New York, NY, USA, 2004. ACM.