

# Block-Based MAP Disparity Estimation Under Alpha-Channel Constraints

Peter J. McNerney, Janusz Konrad, *Senior Member, IEEE*, and Margrit Betke, *Member, IEEE*

**Abstract**—Disparity estimation belongs to the most important, but difficult, problems in image processing and computer vision. Its importance stems from a wide range of applications, while its difficulty is related to ill-posedness. To date, numerous disparity estimation algorithms have been developed. In this paper, we consider a particular case of disparity estimation based on two views and a known alpha channel partitioning each view into foreground and background. The main idea is to use this partitioning in order to enhance disparity estimation in the foreground object close to its boundary. We propose a block-based disparity model with two alpha-channel constraints: a photometric one, disabling invalid intensity/color matches, and a geometric one, preventing disparity smoothing between foreground and background. We incorporate these constraints into a Bayesian framework using the maximum *a posteriori* probability criterion. We experimentally demonstrate improvements in the estimated disparities at foreground object boundaries, and show examples of image relighting using these disparities.

**Index Terms**—Alpha channel, blue screening, disparity estimation, image compositing, stereo.

## I. INTRODUCTION

THE recovery of 3-D scene structure from multiple views, and the related problem of disparity estimation, belong to the most important, but difficult, problems in image processing and computer vision. Their importance stems from a wide range of applications, from 3-D scene analysis, through multiview video coding, to mixing of real and computer-generated 3-D imagery. The difficulties in disparity estimation stem from its ill-posed nature, requiring suitable *a priori* models. The most successful disparity estimation methods to date using such priors are methods based on 1-D correspondences using dynamic programming, 2-D correspondences over rectangular blocks [1], dense vector fields (e.g., Markov random fields) [2], [3], and graph cuts [4]. Many variants of these techniques have been proposed, and are being constantly developed.

In this paper, we consider disparity estimation in the particular context of image compositing and mixed-reality systems. When several images need to be combined together, information about the background and foreground in each image is needed first. This information, known as the alpha channel, can

be acquired by such traditional techniques as “blue screening,” based on color discrimination, or by more recent techniques like “depth keying,” using depth discrimination. In standard image compositing using the “over” method [5], pixel intensities from two or more images are linearly combined based on their opacities (alpha channel values). Since this approach accounts for no object/scene structure, correct 3-D object positioning and accurate lighting are not possible. Therefore, of interest is stereo-based image compositing that exploits 3-D structure for seamless object/scene integration (photometric and geometric).

In order to accurately recover structure of an object captured against blue screen, we propose to exploit alpha channels computed prior to image compositing. The goal is to improve disparity estimates close to foreground object’s boundary. We propose a block-based disparity model with two alpha-channel constraints: a photometric one, disabling invalid intensity/color matches, and a geometric one, preventing disparity (depth) smoothing between foreground and background, and we incorporate both into a Bayesian framework under the maximum *a posteriori* probability criterion. Although pixel transparency has been considered within disparity estimation in the past [6], only a photometric constraint was applied.

The paper is organized as follows. In Section II, we overview image compositing and alpha channel estimation. In Section III, we describe standard block-based disparity estimation, and in Section IV, we extend it by incorporating alpha channel information. We show experimental results in Section V and draw conclusions in Section VI.

## II. IMAGE COMPOSITING USING STEREO

In visual special effects and in mixed-reality systems, several images are often combined together by means of compositing [5]; for example, in the so-called “over” method pixel intensities are linearly combined based on their opacities. Clearly, pixel opacities (visible, partially visible, invisible), also known as alpha channel values, must be known for all images prior to compositing. While for computer-generated images opacities are known, they are unknown for camera-acquired images and need to be computed.

One technique used to separate an object from background in life images is *blue screen matte*; objects are filmed in front of a blue background, after which the blue areas are removed assuming no blue color is present in the foreground object. The main difficulty lies in the accuracy of extracting the matte (alpha channel); fine details, such as hair or fur, can be accidentally cut off from the rest of the figure. Today, with images captured digitally, relatively accurate alpha-channel extraction methods exist [7]. An alternative is *depth keying* [8] where depth of a 3-D scene is captured by special camera, that simultaneously records the photometry and depth, followed by depth processing in order to differentiate between the foreground object (smaller depth) and background (larger depth). This approach, however, requires

Manuscript received March 7, 2006; revised June 9, 2006 and August 11, 2006. This work was supported in part by the Office of Naval Research under Grant N000140110444 and in part by the National Science Foundation under Grants P200A010310, EIA0202067, IIS 039009, 0093367, 0308213.

P. J. McNerney and J. Konrad are with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA (e-mail: jkonrad@bu.edu).

M. Betke is with the Department of Computer Science, Boston University, Boston, MA 02215 USA (e-mail: betke@bu.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSVT.2007.894045

specialized, expensive hardware, and, in some scenarios, is unreliable. Although the depth can be captured implicitly by two (or more) cameras, and then recovered numerically, such techniques have not yet reached performance level of blue-screen methods.

Therefore, in this work, we rely on blue screening followed by Bayesian matte extraction [7]. However, in order to seamlessly insert the captured object into an image of 3-D scene (positioning, relighting), the object's structure (depth) is needed as well. We propose to capture an object in front of blue screen with two video cameras. We focus on depth extraction accounting for object boundaries that can be implemented in real time on today's hardware. Note that when compositing images of a 3-D object and 3-D scene, their illumination parameters need to be considered. While for computer-generated images these parameters are known, they need to be adequately modeled and estimated for camera-acquired imagery [9].

### III. BLOCK-BASED MAP DISPARITY ESTIMATION

We consider a block-based disparity model due to its success in video coding (as motion model), ease of implementation (compared to more advanced disparity models) and available hardware to perform estimation on (application-specific integrated circuits). Consider a pair of images for which we seek disparity: left image  $I_l$  and right image  $I_r$ . Let  $\mathbf{x}$  denote position of a pixel in the domain  $\mathcal{D}$  of either  $I_l$  or  $I_r$ . Let this domain be partitioned into blocks  $\mathcal{B}_k$ , i.e.,  $\mathcal{D} = \cup_{k \in \mathcal{K}} \mathcal{B}_k$ , where  $\mathcal{K}$  denotes the set of indexes for all blocks. Let  $\mathbf{d}_k$  be a disparity vector associated with block  $\mathcal{B}_k$ ; all pixels in  $\mathcal{B}_k$  are assigned the same disparity, very much like motion vectors in MPEG and H.26X video coding standards.

The goal is to compute a disparity field  $\mathbf{d}$ , defined as a field of vectors  $\{\mathbf{d}_k, k \in \mathcal{K}\}$ , that relates *homologous* points in the left and right images, i.e., points that are projections of the same 3-D point in a scene being captured. We use standard observation model relating the images and disparity

$$I_l(\mathbf{x}_i) = I_r(\mathbf{x}_i + \mathbf{d}_k) + \eta_i, \quad \forall \mathbf{x}_i \in \mathcal{B}_k, k \in \mathcal{K} \quad (1)$$

where  $\eta_i$  is a stationary, zero-mean white Gaussian noise term with variance  $\sigma^2$  ( $\eta_i \sim \mathcal{N}(0, \sigma^2)$ ). This model expresses the assumption of intensity constancy between homologous points plus uncertainty due to sensor noise, miscalibration, etc.

The above model is commonly used in video coding with respect to motion. Considered in the context of maximum likelihood (ML) disparity estimation:

$$\hat{\mathbf{d}} = \arg \max_{\mathbf{d} \in \mathcal{S}} p(I_l | \mathbf{d}, I_r) \quad (2)$$

where  $\mathcal{S}$  is the state space of disparity field  $\mathbf{d}$ , and  $p(I_l | \mathbf{d}, I_r)$  is the *likelihood* of obtaining  $I_l$  given both  $I_r$  and disparity field  $\mathbf{d}$ , this model leads to the following disparity block matching:

$$\hat{\mathbf{d}}_k = \arg \min_{\mathbf{d}_k \in \mathcal{S}_k} \sum_{\mathbf{x}_i \in \mathcal{B}_k} (I_l(\mathbf{x}_i) - I_r(\mathbf{x}_i + \mathbf{d}_k))^2, \quad k \in \mathcal{K} \quad (3)$$

where  $\mathcal{S}_k$  is the state space for single disparity vector  $\mathbf{d}_k$ . Note that, due to the implicit independence between neighboring disparity vectors, each vector  $\mathbf{d}_k$  is computed separately.

The ML estimation framework works well in the compression context, whether vis-à-vis motion or disparities. However, our goal is to recover a physical quantity (depth) to be used in a physically-meaningful way (image compositing with relighting). Therefore, we need to account for physical properties of depth, e.g., its local smoothness. In other words, in order to recover smooth disparity (depth) consistent with true object shape we need to impose an additional constraint.

We accomplish this through Bayesian framework that proved successful in this context in the past [2], [3], [10]. We formulate the disparity estimation as a maximum *a posteriori* probability (MAP) problem

$$\begin{aligned} \hat{\mathbf{d}} &= \arg \max_{\mathbf{d} \in \mathcal{S}} p(\mathbf{d} | I_l, I_r) = \arg \max_{\mathbf{d} \in \mathcal{S}} \frac{p(I_l | I_r, \mathbf{d}) p(\mathbf{d} | I_r) p(I_r)}{p(I_l, I_r)} \\ &= \arg \max_{\mathbf{d} \in \mathcal{S}} p(I_l | I_r, \mathbf{d}) p(\mathbf{d} | I_r) \quad (4) \end{aligned}$$

where  $p(\mathbf{d} | I_l, I_r)$  is the *a posteriori* probability of disparity field  $\mathbf{d}$  given the two images. The prior probability  $p(\mathbf{d} | I_r)$  expresses assumptions that we make about the disparity field  $\mathbf{d}$ . We assume that  $\mathbf{d}$  is independent of  $I_r$  ( $p(\mathbf{d} | I_r) = p(\mathbf{d})$ ), i.e., that the knowledge of the right image tells us nothing about the disparity field itself. Although this assumption does not hold everywhere in the image (the knowledge of object boundaries in the right image may suggest the location of discontinuities in the disparity/motion field [11]), it is nevertheless widely, and successfully, used in practice. We model the disparities  $\mathbf{d}$  using a vector Markov random field (MRF), so that  $p(\mathbf{d})$  becomes a Gibbs distribution. For a detailed development of scalar MRF models for images, we refer the reader to the seminal work of Geman and Geman [12], and of vector MRF models for motion—to the works of Murray and Buxton [13], and of Konrad and Dubois [11].

Since the likelihood and prior distributions are exponential, the maximization (4) can be rewritten as energy minimization

$$\hat{\mathbf{d}} = \arg \min_{\mathbf{d} \in \mathcal{S}} \sum_{k \in \mathcal{K}} \left[ \sum_{\mathbf{x}_i \in \mathcal{B}_k} (I_l(\mathbf{x}_i) - I_r(\mathbf{x}_i + \mathbf{d}_k))^2 + \lambda \sum_{l \in \mathcal{N}_k} \|\mathbf{d}_k - \mathbf{d}_l\|^2 \right] \quad (5)$$

where  $\mathcal{N}_k$  is the block-neighborhood of block number  $k$ , i.e.,  $l \in \mathcal{N}_k$  means that block number  $l$  is a neighbor of block number  $k$ . We use the first-order neighborhood, i.e., top, bottom, left, and right block neighbors.

The summation over block index  $k$  accumulates energy over all blocks. The energy of block number  $k$  consists of the data matching error (summation over all pixels in  $\mathcal{B}_k$ ) plus the disparity smoothness error (summation over neighboring-block index  $l$ ). Note that because of the latter term, unlike in the ML case, the individual disparities  $\mathbf{d}_k$  cannot be estimated independently, and an iterative solution algorithm is required. The balance between intensity matching and disparity smoothness errors is controlled by the weight  $\lambda$  which is proportional to the noise variance  $\sigma^2$  from the observation model (1); the larger this variance, the larger the  $\lambda$ , and thus the more confidence is given to disparity smoothness.

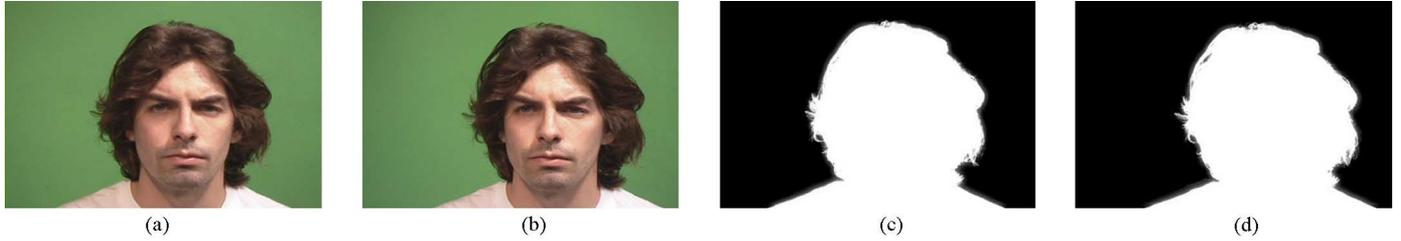


Fig. 1. Original (a) left and (b) right images captured against green screen. (c) Left and (d) right alpha channels computed by method of Chuang *et al.* [7].

We solve the ML minimization (3) by a single-pass exhaustive search over  $\mathcal{S}$  with specified accuracy (e.g., quarter-pixel). In the case of MAP minimization (5), we apply exhaustive search iteratively until the overall energy change is sufficiently small or the maximum number of iterations is reached.

#### IV. BLOCK-BASED MAP DISPARITY ESTIMATION UNDER ALPHA-CHANNEL CONSTRAINTS

By the very nature of blue-screen object acquisition, true depth discontinuities at the object boundary are associated with color discontinuities (captured by the alpha channel). Standard block-based disparity estimation methods, described in the previous section, are not designed to handle such discontinuities. We propose an extension of block-based MAP disparity estimation to permit disparity (depth) discontinuity based on the known alpha channels. The proposed models could equally apply to segmentation of the foreground object itself in order to permit in-object depth discontinuities (e.g., implicit [10] or explicit [3]), however in this work we assume the captured object is internally smooth (e.g., face).

In the MAP formulation (5), the disparity estimation algorithm is applied uniformly to all blocks in the reference image (left image in our case). However, because of the intended image compositing, there is no need to estimate disparities in the constant-color background (only depth of the visible parts of the object, e.g., face, is needed). This permits us to modify the domain of the disparity search  $\mathcal{K}$  to exclude those blocks that belong to the background. Let  $\mathcal{K}_{\alpha_l}$  be the set of block indexes such that for  $k \in \mathcal{K}_{\alpha_l}$ , at least one pixel from  $\mathcal{B}_k$  belongs to the left-image alpha plane  $\alpha_l$  (foreground object in the left image). This new restricted domain results solely in speeding up the disparity estimation process.

The alpha channel can be also used to eliminate erroneous matches. Since we assumed that color of the background screen is selected so that it is absent from the object, we know that a match between an object block and a block from the constant-color background is very unlikely. Thus, we can restrict the state space of each disparity vector by incorporating the alpha channel  $\alpha_r$  of the right image. Since the new state space is now spatially variant, its practical implementation is more complex. We handle this by introducing a gain factor  $\rho(\mathbf{x})$  dependent on the right alpha channel

$$\rho(\mathbf{x}) = \begin{cases} 1, & \text{for } \alpha_r(\mathbf{x}) > 0 \\ \infty, & \text{for } \alpha_r(\mathbf{x}) = 0. \end{cases}$$

Clearly, for a pixel at  $\mathbf{x}$  within the object ( $\alpha_r(\mathbf{x}) > 0$ ), the gain  $\rho$  is equal to 1, while for a pixel belonging to the back-

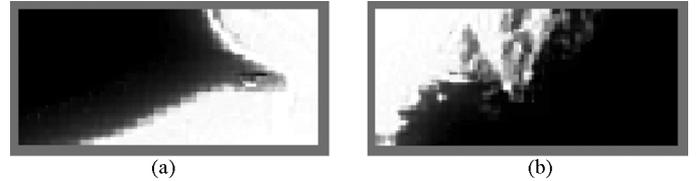


Fig. 2. Closeups from: (a) left and (b) right alpha channels from Fig. 1(c) and (d), respectively, showing fuzzy transitions.

ground it is infinite (in practice, very large). If the matching error  $(I_l(\mathbf{x}) - I_r(\mathbf{x} + \mathbf{d}_k))^2$  in (5) were multiplied by  $\rho(\mathbf{x} + \mathbf{d}_k)$ , then a left-image block  $\mathcal{B}_k$  entirely within the foreground object would likely find a match within the right-image foreground; otherwise the gain  $\rho(\mathbf{x} + \mathbf{d}_k)$  would lead to large energy values. Blocks at the object boundary would also likely find a match within the object or at its periphery, but not in the background. The introduced multiplicative gain leads to a *photometric constraint* on possible solutions, and is expected to result in fewer false matches at object boundaries.

Note that the smoothness constraint has been applied so far uniformly throughout the disparity field, also between blocks of which one belongs to the object while the other does not. Since such blocks are positioned across an object boundary, the smoothness constraint should be disabled. This can be achieved by using object boundary in the left image *via* the left alpha channel  $\alpha_l$ . Let  $\gamma_k$  be an indicator function defined for the left alpha channel as follows:

$$\gamma_k = \begin{cases} 1, & \text{if } \exists \mathbf{x} \in \mathcal{B}_k \text{ s.t. } \alpha_l(\mathbf{x}) > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Clearly,  $\gamma_k$  equals 1 for blocks at least partially overlapping the object (foreground) area of the left image. For blocks entirely in the background,  $\gamma_k = 0$ . If  $\delta[\cdot]$  is the Kronecker delta ( $\delta[n]$  equals 1 for  $n = 0$ , and 0 otherwise), then for blocks number  $k$  and  $l$ ,  $\delta[\gamma_k - \gamma_l] = 1$  only when both of them belong to either the foreground or background. Multiplying the prior term  $\|\mathbf{d}_k - \mathbf{d}_l\|^2$  in (5) by  $\delta[\gamma_k - \gamma_l]$ , leads to a *geometric constraint* disabling disparity smoothness across object boundary.

With the above improvements, the new MAP disparity estimation can be expressed as follows:

$$\hat{\mathbf{d}} = \arg \min_{\mathbf{d} \in \mathcal{S}} \sum_{k \in \mathcal{K}_{\alpha_l}} \left[ \sum_{\mathbf{x}_i \in \mathcal{B}_k} \rho(\mathbf{x}_i + \mathbf{d}_k) (I_l(\mathbf{x}_i) - I_r(\mathbf{x}_i + \mathbf{d}_k))^2 + \lambda \sum_{l \in \mathcal{N}_k} \delta[\gamma_k - \gamma_l] \|\mathbf{d}_k - \mathbf{d}_l\|^2 \right]. \quad (6)$$

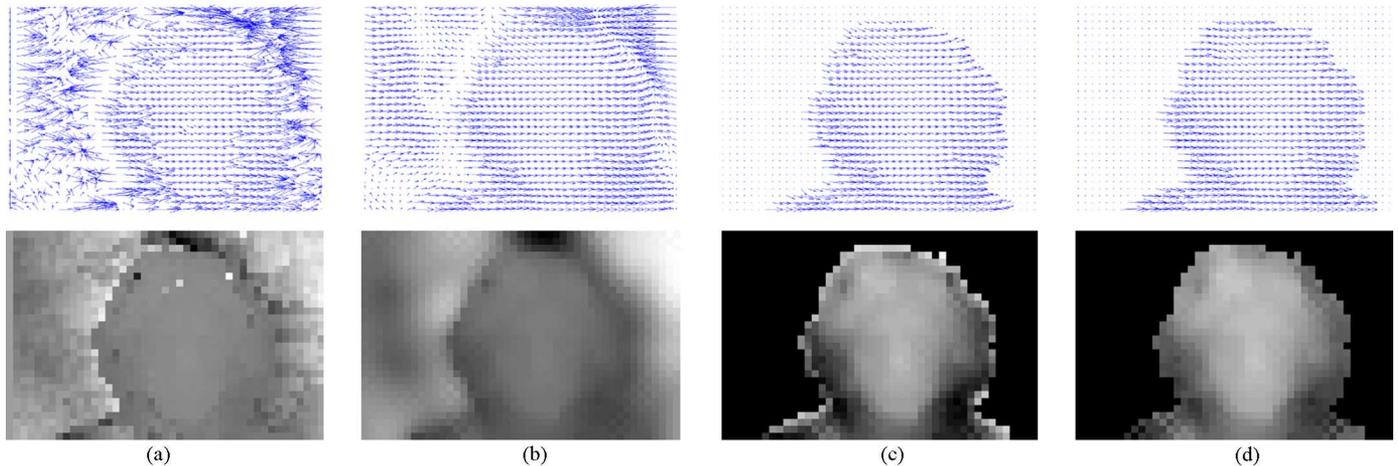


Fig. 3. Disparity estimates for the test image from Fig. 1. (a) ML; (b) MAP; (c) MAP with photometric alpha-channel constraint. (d) MAP with photometric and geometric alpha-channel constraints, all shown as vector field and as grayscale of the horizontal disparity [in (c) and (d), background disparity is not estimated and set arbitrarily to zero/black].

Note the  $\alpha$ -induced restrictions implemented through the new search domain  $\mathcal{K}_{\alpha_i}$ , gain  $\rho(\mathbf{x}_i + \mathbf{d}_k)$  and indicator function  $\delta[\gamma_k - \gamma_i]$ . The above minimization is accomplished again iteratively by exhaustive search.

## V. EXPERIMENTAL RESULTS

Fig. 1 shows a typical head-and-shoulders stereo pair used in our experiments along with corresponding alpha channels. The stereo pair was captured by an uncalibrated, roughly parallel camera setup. This is a common scenario in stereo acquisition for entertainment purposes; camera calibration [14] would help achieve more accurate results. The alpha channel was computed by first thresholding the red, green and blue image components at each pixel. Since uneven lighting, shadows, and fuzzy areas contribute to alpha-channel errors, we refined the initial estimate by applying a Bayesian approach proposed by Chuang *et al.* [7]. That approach uses a spatial model for the alpha channel within a maximum *a posteriori* probability criterion. Note the fuzzy transition between the foreground (white) and background (black) that permits a seamless composition of object with another background (Fig. 2).

In order to recover disparities, we first applied the ML estimation (3), which basically implements standard block matching on disparities. Since only the intensity matching error is minimized ( $\lambda = 0$ ), disparity estimates should be little correlated spatially (large spatial variability) similarly to fast [15] or multiresolution [16] block matching implementations that also include no prior vector model. This is confirmed in Fig. 3(a) that shows ML disparity estimate as a vector plot and as a brightness image of the horizontal disparity component. Note significant disparity variability in the background and at the object boundary. Subsequently, we used the ML disparity estimate as an initial solution for the MAP estimation (5); results are shown in Fig. 3(b). Note that, as expected, the MAP estimate is much less noisy (more spatially consistent) than the ML estimate, due to the added disparity smoothness constraint. It is more consistent locally and better reflects the actual scene structure (smooth

object surface, fairly constant depth in the background). However, a transition band around object boundary is formed by the vectors, where a gradual (smooth) change of disparity vector length and orientation takes place between the foreground and background.

Fig. 3(c) shows a disparity estimated using the new  $\alpha$ -constrained algorithm but only with the photometric constraint (gain  $\rho$  penalizing foreground-background matches). The background disparities were not computed, but instead were arbitrarily set to zero (they are not needed in image compositing). In the foreground object, the transition-band effect has been largely eliminated and a clear disparity discontinuity has formed. Also, many false matches in the foreground have been eliminated. Although the vector field has been improved compared to the standard MAP case Fig. 3(b), one can still notice disparity variations along object boundary (darker and brighter blocks on the boundary) predominantly due to the isotropic smoothness constraint. In order to correct this, the disparity smoothness term needs to be disabled across object boundary. Fig. 3(d) shows a disparity estimate obtained by the proposed MAP algorithm with both alpha-channel constraints turned on (6). Clearly, the disparity field is very consistent along object boundary and the overall appearance of depth is more smooth accurately depicting the face structure. This improvement is further evidenced by the difference field of horizontal disparities shown in Fig. 4(a); disparities get corrected only at object boundary and are unchanged inside.

Since the intended application of the proposed algorithm is stereo-based image compositing and relighting, we present an example in this context. Note that the estimated alpha-constrained disparity field, although smooth, is still block-constant. Clearly, the corresponding depth, computed from the horizontal disparity component assuming known camera baseline and focal length, is block-constant as well. Although dense disparity field could have been computed (at high computational cost), we found that very good results can be obtained by smoothing block-constant disparity with Gaussian filter. A Gaussian-smoothed depth is shown in Fig. 4(b); the result is

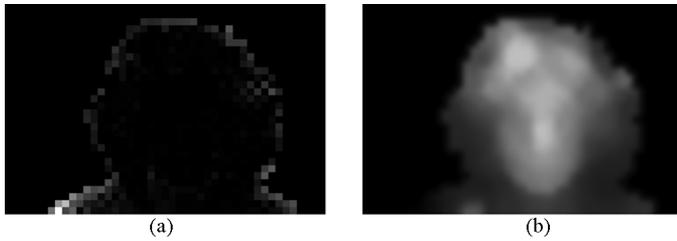


Fig. 4. (a) Absolute difference of horizontal disparity for the results from Fig. 3(c) and (d) shown as brightness image (black—no difference); and (b) final depth field after Gaussian filtering.



Fig. 5. Examples of two different illuminations of the foreground object.

smooth but with clearly protruding nose, chin, and some hair, and retracted eye sockets, lower hair, shoulders.

Based on the estimated depth and intensity of the original stereo pair, we computed a textured mesh model using commercial program *Truespace* [17]. As object surface texture, we used intensity and color of the left image, and as its transparency map—the left alpha channel. Once the object's shape was reconstructed, we applied a new illumination and rendered a new image, again using *Truespace*.

Fig. 5 shows a rendering of the head from our test image [Fig. 1(a) and (b)] with a moving light source. Note that the light bends around the face of the person (especially around the nose), as it would if there were an actual physical light in the scene. Clearly, the estimated disparity (depth) helps achieve plausible illumination effects that would have been impossible using one image only (no stereo).

## VI. SUMMARY AND CONCLUSIONS

We proposed an improved block-based MAP disparity estimation that accounts for alpha-channel boundaries. The resulting disparities exhibit more consistent behavior along foreground object boundaries, and lead to smoother depth estimates in the foreground than standard block-based ML and MAP techniques. Since the proposed algorithm performs at most 5 iterations of calculations of complexity similar to those used in block matching, a real-time hardware implementation is foreseeable today, especially with fast algorithms [15]. The obtained disparity estimates were used for foreground object relighting

demonstrating feasibility of image compositing/relighting using stereo. The current algorithm is applied independently to consecutive frame pairs of stereoscopic video thus likely leading to inconsistent depth estimates in time. One possible extension of the proposed algorithm could be to assure temporal continuity of depth by means of an additional energy term. Also, disparity discontinuities are currently permitted only at discontinuities of the alpha channel. However, foreground objects themselves may include depth discontinuities; one could use a “weak membrane” [18] or line process [11] model to permit such depth discontinuities within the foreground object.

## REFERENCES

- [1] T. Kanade and M. Okutomi, “A stereo matching algorithm with an adaptive window: Theory and experiment,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 9, pp. 920–932, Sep. 1994.
- [2] R. March, “Computation of stereo disparity using regularization,” *Pattern Recognit. Lett.*, vol. 8, pp. 181–187, Oct. 1988.
- [3] L. Alvarez, R. Deriche, J. Sanchez, and J. Weickert, “Dense disparity map estimation respecting image discontinuities: A PDE and scale-space based approach,” *J. Vis. Commun. Image Represent.*, vol. 13, pp. 3–21, Mar. 2002.
- [4] O. Boykov, O. Veksler, and R. Zabih, “Fast approximate energy minimization via graph cuts,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1231, 2001.
- [5] T. Porter and T. Duff, “Compositing digital images,” in *Proc. SIGGRAPH'84*, Jul. 1984, vol. 18, pp. 253–259.
- [6] R. Szeliski and P. Golland, “Stereo matching with transparency and matting,” *Int. J. Comput. Vis.*, vol. 32, no. 1, pp. 45–61, 1999.
- [7] Y. Chuang, B. Curless, D. Salesin, and R. Szeliski, “A Bayesian approach to digital matting,” in *Proc. IEEE Conf. Comp. Vis. Pattern Recognit.*, Dec. 2001, vol. 2, pp. 264–271.
- [8] R. Gvili, A. Kaplan, E. Ofex, and G. Yahav, “Depth keying,” in *Proc. SPIE Stereoscopic Displ. Virtual Reality Syst.*, May 2003, vol. 5006, pp. 564–574.
- [9] B. Horn, *Robot Vision*. Cambridge: MIT Press, 1986.
- [10] A. Mancini and J. Konrad, “Robust quadtree-based disparity estimation for the reconstruction of intermediate stereoscopic images,” in *Proc. SPIE Stereoscopic Displ. Virtual Reality Syst.*, Jan. 1998, vol. 3295, pp. 53–64.
- [11] J. Konrad and E. Dubois, “Bayesian estimation of motion vector fields,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 9, pp. 910–927, Sep. 1992.
- [12] S. Geman and D. Geman, “Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, Nov. 1984.
- [13] D. Murray and B. Buxton, “Scene segmentation from visual motion using global optimization,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 2, pp. 220–228, Mar. 1987.
- [14] F. Pedersini, A. Sarti, and S. Tubaro, “Multi-camera systems: Calibration and applications,” *IEEE Signal Process. Mag.*, vol. 16, no. 3, pp. 55–65, May 1999.
- [15] V.-A. Nguyen and Y.-P. Tan, “Efficient block-matching motion estimation based on integral frame attributes,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 3, pp. 375–385, Mar. 2006.
- [16] J. Zan, M. Ahmad, and M. Swamy, “Comparison of wavelets for multiresolution motion estimation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 3, pp. 439–446, Mar. 2006.
- [17] Caligari Corporation, *Truespace 4.0/6.5*. 2006 [Online]. Available: [www.caligari.com](http://www.caligari.com)
- [18] M. Figueiredo and J. Leitão, “Simulated tearing: An algorithm for discontinuity-preserving visual surface reconstruction,” in *Proc. IEEE Conf. Comp. Vis. Pattern Recognit.*, Jun. 1993, pp. 28–33.