SPATIO-TEMPORAL BINARY VIDEO INPAINTING VIA THRESHOLD DYNAMICS

M. Oliver *R.P. Palomares* *C. Ballester* *G. Haro**

DTIC – Universitat Pompeu Fabra

ABSTRACT

We propose a new variational method for the completion of moving shapes through binary video inpainting that works by smoothly recovering the objects into an inpainting hole. We solve it by a simple dynamic shape analysis algorithm based on threshold dynamics. The model takes into account the optical flow and motion occlusions. The resulting inpainting algorithm diffuses the available information along the space and the visible trajectories of the pixels in time. We show its performance with examples from the Sintel dataset, which contains complex object motion and occlusions.

***Index Terms*—** Shape completion, binary video inpainting, threshold dynamics.

1. INTRODUCTION

Video inpainting stands for the completion of missing, damaged or occluded information in a video sequence in such a way that this restoration is as unnoticeable (visually plausible) as possible. The applications include tools for cinema post-production to remove, e.g., unwanted or private items, or for the recovering of occluded areas in new-view generation for 3D television or broadcasting of sport events, to mention just a few. First attempts on video inpainting consisted in applying image inpainting techniques to each frame separately but the temporal incoherence from frame to frame is very noticeable for the human vision system, producing an undesirable flickering effect. Video inpainting brings additional challenges to the ones of image inpainting not only in order to obtain temporally coherent results but also due to the occlusions and disocclusions between objects that move along time.

Binary inpainting methods aim at recovering (or disoccluding) shapes, stated as binary objects. They can be a tool for the automatic understanding of a dynamic scene through its decomposition in completed and isolated objects interacting among them. On the other hand, they might be combined with texture-based inpainting, in a two-step algorithm, so that the completed shape helps to guide the copy of patches: inside the shape of interest only patches from the same object are allowed to be copied and similarly for the background.

This paper proposes a binary video inpainting method that works directly in the spatio-temporal dimension. To the best of our knowledge, this is the first work on binary video inpainting, that is the completion of moving shapes. We propose a variational formulation for object-based video inpainting that recovers a smooth surface by imposing not only spatial regularity but also temporal continuity along the visible trajectory of the object. It includes the convective derivative and has no restrictions on background nor foreground movements. The convective derivative has been used for ensuring spatio-temporal consistency in other video editing applications [6, 5, 17, 31].

Trajectories and the convective derivative are defined by the optical flow (the vector field that recovers the apparent motion of two consecutive frames) which is previously estimated and completed (inside the inpainting mask or hole) also through variational methods. and we present qualitative and quantitative results showing that our binary video inpainting method obtains similar results using as input either the ground truth optical flow or an estimated one. On the other hand, the optical flow is unknown inside the hole and it is interpolated with a motion inpainting method.

One of the main difficulties that has to be tackled in video completion is due to occlusion effects. Object occlusions and disocclusions generate artifacts which are specially visible at moving occlusion boundaries. Moreover, optical flow methods may fail in occlusion areas due to unreliable shape or point matching. In general, points visible at time t that are occluded at time $t+1$ should not have a corresponding point at frame $t+1$. Thus video completion algorithms have to detect such occlusions in order to correctly decide how to interpolate. Our method keeps track of the motion occlusions, which are estimated from the optical flow, and incorporates them into the proposed variational method.

Most inpainting approaches may be divided into geometry-oriented methods (e.g., [23, 2, 10, 22, 11]), texture-oriented methods (e.g., [12, 36, 28, 1, 7, 18, 26, 33, 37, 13]) and methods combining both ideas. The work of Cao et al. [9] is an example of the later and combines an exemplar-based approach with a geometric guide computed by minimizing Euler's elastica of contrasted level lines in the inpainted region. On the other hand, binary inpainting tools for images are also used to disocclude shapes and thus can be considered as geometry-oriented methods [14, 4, 27]. Merriman, Bence, and Osher argue in [25] that the convolution of the indicator function of a shape with a Gaussian followed by a threshold at $1/2$ simulates the mean curvature motion. This was proved by Barles and Georgelin [3] and Evans [16]. The diffusion process followed by thresholding is known as threshold dynamics and provides a dynamic shape analysis. Threshold dynamics interpolations usually minimize a geometric functional, based either on the length, area, or curvature of the shape contours [25, 19] and some authors have used it for 2D-shape completion or disocclusion [14, 4, 27].

2. PROPOSED MODEL

We present in this section a variational model for video inpainting with a threshold dynamics strategy using a differential operator based on a generalized 3D gradient with the convective derivative in time and the usual gradient in space.

Let $u_0(\mathbf{x}, t)$ be a binary video sequence defined on $\mathcal{V}\backslash\mathcal{M}$, where $\mathcal{V} = \{(\mathbf{x}, t) : \mathbf{x} = (x, y) \in \Omega,\ t \in \mathbb{R}\}$ and $\mathcal{M} \subset \mathcal{V}$ denotes the inpainting hole with missing information. Here, $\Omega \subset \mathbb{R}^2$ is assumed to be the image domain (i.e., the spatial domain of any image frame

*The authors acknowledge partial support by MICINN project, reference MTM2012-30772, by TIN2015-70410-C2-1-R (MINECO/FEDER, UE) and by GRC reference 2014 SGR 1301, Generalitat de Catalunya.

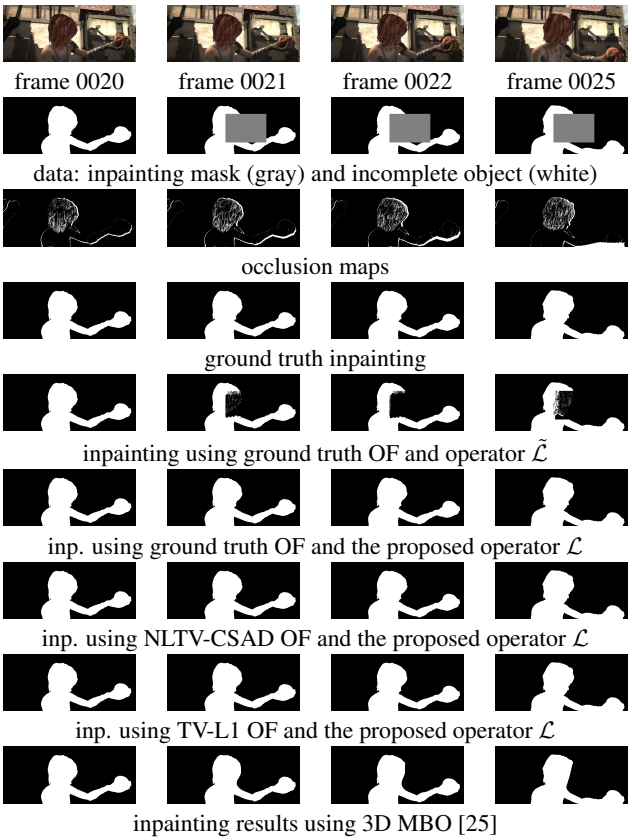

Fig. 1: Experiment with alley_1 sequence from Sintel [8]: Inpainting results with different methods and optical flow estimations.

at time t) which is, as usual, a rectangle in $\mathbb{R}^2$.

Let us assume that we have at our disposal an optical flow method as well as an optical flow inpainting method allowing to compute the optical flow $\mathbf{v}(\mathbf{x}, t)$ for all $(\mathbf{x}, t) \in \mathcal{V}$. Later on we will precise the proposed estimation of $\mathbf{v}$ in $\mathcal{V}$.

In order to inpaint the binary video inside the inpainting mask $\mathcal{M} \subset \mathcal{V}$ we propose to solve the following optimization problem

$$\min_{u:\mathcal{V}\to\{0,1\}} \int_{\mathcal{M}} \|\mathcal{L}(u)\|^2, \quad \text{s.t. } u = u_0 \text{ in } \mathcal{V} \setminus \mathcal{M} \tag{1}$$

where $\mathcal{L}(u)$ is the differential operator defined taking into account both spatial and temporal regularity as well as the occlusion areas produced by the motion of objects in the scene, that is:

$$\mathcal{L}(u) = (u_x, u_y, \gamma\chi\partial_{\mathbf{v}} u), \tag{2}$$

where $\gamma > 0$ is a parameter and $\chi : \mathcal{V} \to [0, 1]$ is a function modeling the occlusion areas so that $\chi(\mathbf{x}, t) = 0$ identifies the occluded pixels, i.e. pixels that are visible at time t but not at time $t+1$. Thus, $\chi(\mathbf{x}, t) = 1$ identifies the non occluded pixels and the functional only imposes temporal regularity along the pixel trajectories that are not occluded. The convective derivative is defined as

$$\partial_{\mathbf{v}} u(\mathbf{x}, t) = \nabla u(\mathbf{x}, t) \cdot \mathbf{v}(\mathbf{x}, t) + \frac{\partial u}{\partial t}(\mathbf{x}, t). \tag{3}$$

Let us recall how it naturally appears: From the assumption that for a Lambertian object under uniform constant illumination, the brightness of an object's particle does not change in time, one deduces that $u(\mathbf{x}(t), t)$ is constant along trajectories of the points in the scene. This implies that

$$0 = \frac{du}{dt}(\mathbf{x}(t), t) = \nabla u(\mathbf{x}, t) \cdot \frac{d\mathbf{x}(t)}{dt} + \frac{\partial u}{\partial t}(\mathbf{x}, t) \approx \partial_{\mathbf{v}} u(\mathbf{x}, t), \tag{4}$$

which leads to the well-known brightness constancy assumption [20] since $\mathbf{v} \approx d\mathbf{x}(t)/dt$. In our case, since u represents a binary function that identifies a shape, by minimizing (1), we are imposing shape regularity along the trajectories (thanks to the convective derivative) and also spatial smoothness in the recovered shape (thanks to the spacial derivatives in the operator (2)). Moreover, since we do not consider the convective derivative for occluded pixels, see eq. (2), we are imposing the regularity only along the visible trajectories.

In our proposal (1) – (2), the parameter γ accounts for the different units in the spatial and temporal domains and also balances the effect of the temporal diffusion in the resulting gradient-descent equation. Observe that, when γ is big enough, the minimization of (1) could be approximated by using, instead of $\mathcal{L}$, the operator

$$\tilde{\mathcal{L}}(u) = \partial_{\mathbf{v}} u \tag{5}$$

as the spatial derivatives have almost no impact. However, as the experiments in Sect. 4 show, it is necessary to consider the dynamic evolution of the 3D shape (both in space and time) to correctly complete the moving objects due to the fact that time diffusion is not able to deal with occlusions. In these situations the spactial diffusion helps to complete the shape. An experimental comparison of using $\tilde{\mathcal{L}}$ instead of our proposed $\mathcal{L}$ is shown in Fig. 1 and Fig. 2.

2.1. Optical Flow estimation

To fully specify our method, one needs to provide an estimation of the optical flow. We propose to use, on the original sequence, the optical flow variational method proposed in [30], which can be applied to any energy. In particular, we apply it to both the well-known TV-L1 energy functional [38] and the NLTV-CSAD energy functional that uses Non Local Total Variation as regularization term [35] and a smooth variant of the Census Transform [34] as data term. To show the robustness of our binary inpainting method, we present in Fig. 1 and 2 experiments showing that similar results are obtained using as input either the ground truth optical flow or the estimated ones.

Moreover, as a consequence of removing objects from the sequence, we need to modify the optical flow in the area of the removed object (which constitutes the inpainting mask or hole). So, we need to do motion inpainting inside the hole. We propose to use the optical flow inpainting method proposed in [29], although other methods exist in the literature [21, 24]. In order to correctly fill in the optical flow we dilate the hole due to the irregularities of the optical flow close to its boundary. Usually, the optical flow estimated by variational methods is not accurate at motion boundaries.

2.2. Occlusion estimation

In order to estimate motion occlusions we stem from the assumption that the occluded region, given in our context by $\chi(x, y) = 0$, may be correlated with the region where the divergence of the optical flow is negative. This was pointed out by Sand and Teller in [32], who noticed that the divergence of the motion field may be used to distinguish between different types of motion areas. Schematically, the divergence of a flow field is negative for occluded areas, positive for disoccluded, and near zero for the matched areas. In our method,

we relax this criteria and use the estimation of occluded ($\chi = 0$) and visible ($\chi = 1$) regions as

$$\chi(\mathbf{x}, t) = \begin{cases} 1 & \text{div}(\mathbf{v}) \geq -0.5, \\ 0 & \text{else.} \end{cases} \tag{6}$$

3. ALGORITHM

Following the idea of [15] we propose to modify the original minimization problem (1) by not restricting the solution to be binary and using instead a double well potential in the functional. Then, the minimization problem we consider is:

$$\min_u \int_{\mathcal{M}} \varepsilon \|\mathcal{L}(u)\|^2 + \frac{1}{\varepsilon} W(u), \quad \text{s.t. } u = u_0 \text{ in } \mathcal{V} \setminus \mathcal{M}. \tag{7}$$

where $\varepsilon > 0$ and $W : \mathbb{R} \to \mathbb{R}$ is a double well potential with equidepths at 0 and 1, namely $W(u) = u^2(1-u^2)$. The gradient descent equation for the above functional is:

$$u_s = 2\varepsilon \left(\Delta u + \gamma^2 (\chi \partial_{\mathbf{v}})^* \chi \partial_{\mathbf{v}} u \right) - \frac{1}{\varepsilon} W'(u), \tag{8}$$

where $(\chi \partial_{\mathbf{v}})^*$ denotes the adjoint operator of $\chi \partial_{\mathbf{v}}$.

As in [15], we propose to solve the boundary value problem associated to the PDE (8) by time splitting in such a way that one of the resulting equations, $u_s = -\frac{1}{\varepsilon} W'(u)$, is an ordinary differential equation that is solved by a thresholding step, as in the MBO scheme. Then, starting by an initial spatio-temporal shape $\mathcal{T}^0$ and, considering its (binary) characteristic function $u^0 = \mathbb{1}_{\mathcal{T}^0}$, the core of the threshold dynamics scheme that we propose consists of the iteration of the following steps until convergence:

1. Diffusion step. Compute $\bar{u}(\tau)$, the solution of the following PDE for a certain small diffusion time τ, with initial condition $\bar{u}(0) = \mathbb{1}_{\mathcal{T}^n}$.

$$u_s = \Delta u + \gamma^2 (\chi \partial_{\mathbf{v}})^* \chi \partial_{\mathbf{v}} u$$

2. Thresholding step. Binarize by defining the shape $\mathcal{T} = \{\mathbf{x} : \bar{u}(\tau)(\mathbf{x}) \geq \frac{1}{2}\}$
3. Fidelity step. $\mathcal{T}^{n+1} = (\mathcal{T} \cap \mathcal{M}) \cup \left(\mathcal{T}^0 \cap (\mathcal{V} \setminus \mathcal{M})\right)$

The third step imposes that the binary video coincides with the original video outside the inpainting domain, as done in [14] for binary image inpainting by threshold dynamics.

4. RESULTS

In this section we provide some results of the proposed method used on some image sequences from the Sintel database [8]. We present two types of experiments. First, we consider an inpainting mask that covers part of an object and we apply our proposed method to fill in the object of interest. This situation appears when one needs to recover damaged videos. In order to test our method, we built some synthetic examples where we know the inpainting ground truth. These experiments will help us to evaluate how sensitive our method is to the given optical flow, to set the parameters, and also to compare with the 3D MBO suggested in [25] that evolves a surface by mean curvature motion. In the second type of experiments, the goal is to remove an object from the input video which is occluding another one and we apply the inpainting to complete the occluded object. As previously said in Sect. 2 our model only has two parameters to fix: γ, the balance between the spatial and temporal derivatives, and τ, the diffusion time. We experimentally observed that γ has to be at least 1 to obtain good results. So, we fix it to $\gamma = 1.5$ for all the experiments. Regarding the diffusion time, τ, it is well acknowledged from the threshold dynamics methods that τ has to be big enough to allow for the curve to evolve, but small enough so that the MBO scheme approximates motion by mean curvature as the solution of the Allen-Cahn equation. In all our experiments we set $\tau = 1$.

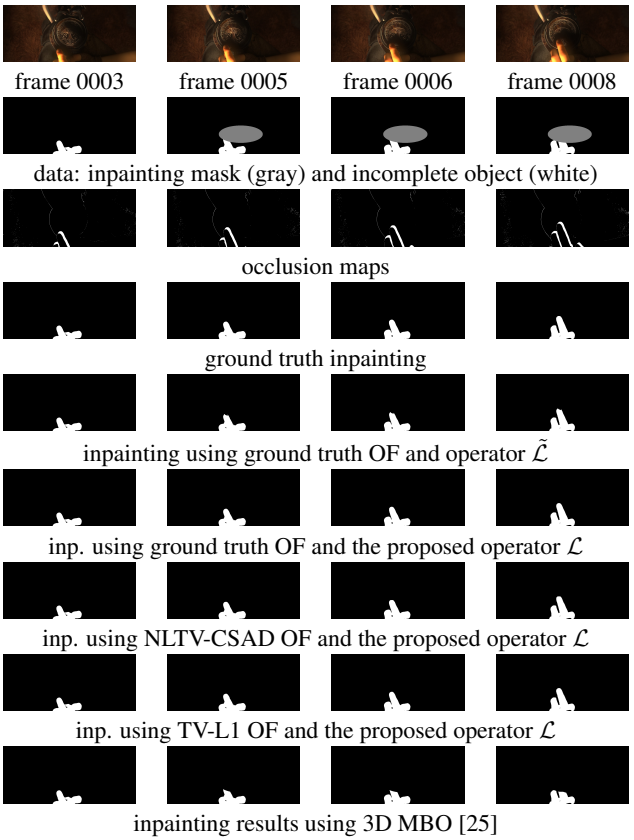

Fig. 2: Experiment with shaman_3 sequence from Sintel [8]: Inpainting results with different methods and optical flow estimations.

Experiments where a damaged object is recovered

The experiments shown in Figs. 1 and 2 illustrate the behaviour of our video inpainting method with the proposed operator $\mathcal{L}$, the operator $\tilde{\mathcal{L}}$, and a comparison with the 3D MBO method. The input video of Fig. 1 is formed by 7 consecutive frames (20 to 26) from the alley_1 sequence, and the video of Fig. 2 is formed by 7 consecutive frames (3 to 9) from the shaman_3 sequence. The first row of both figures displays some of them. In the second row, the object to be inpainted is shown in white and the occluded zone is in gray. As it can be seen in the second row of Figs. 1 and 2, the first frame does not have any pixel occluded. This is because in these experiments we consider the first and last frame to be completely known, i.e, the object to be completed is fully visible in these two frames. The third row contains the ground truth occlusions and the fourth one the ground truth of the object completion. The fifth row shows the result of the inpainting when we use the operator $\tilde{\mathcal{L}}$ and the sixth row corresponds to the inpainting using operator $\mathcal{L}$, both using the

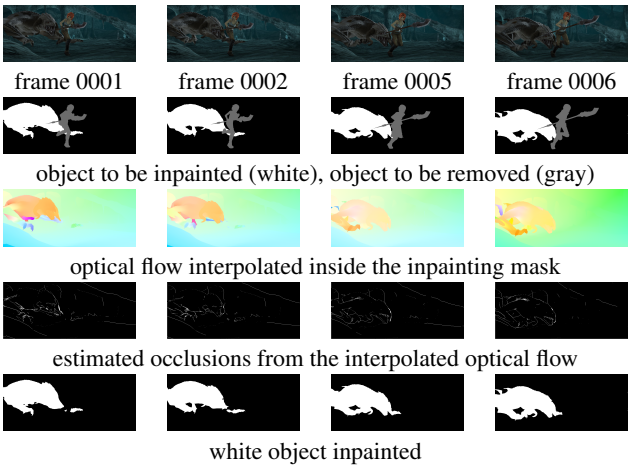

Fig. 3: Removal of an object in a video sequence (cave2).

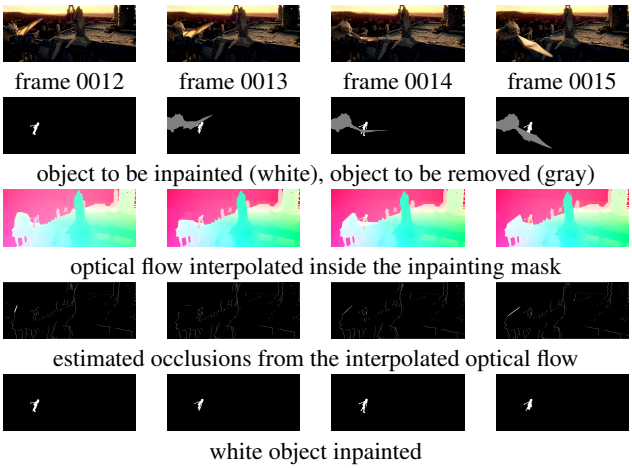

Fig. 4: Removal of an object in a video sequence (temple2).

optical flow ground truth. Both figures show how crucial can be to use the $\mathcal{L}$ operator: In Fig. 1 the hair of the girl is not completely recovered if we only consider the convective term, and in Fig. 2 the top of the finger is incomplete. The reason is that the pixels that need to be inpainted have an occluded trajectory – as it can be seen in the respective occlusion maps (third row) – and no temporal diffusion is applied on them. The spatial diffusion helps to complete these occluded areas. Finally the last row shows the MBO method [25] extended to 2D+time considered as 3D data. As it can be seen in the figures this result does not follow the trajectory of the moving objects but completes according to the smoothness of the shapes: In the last frame of Fig. 1 the face is not well recovered and in Fig. 2 the fingers are cut (completed by a plane in 2D+time), while in both cases our operator $\mathcal{L}$ correctly completes the shapes. For a quantitative evaluation, we present in Table 1 the root mean squared error of three methods: our proposal (1) – (2), using $\tilde{\mathcal{L}}$ instead of $\mathcal{L}$, and MBO. Our operator $\mathcal{L}$ always provides a smaller error.

	MBO [25]	$\tilde{\mathcal{L}}$	$\mathcal{L}$
alley_1	0.18	0.55	0.06
ambush_4	0.46	0.54	0.26
market_5	0.34	0.23	0.07
shaman_3 (seq.1)	0.25	0.10	0.05
shaman_3 (seq.2)	0.63	0.63	0.48
temple_3	0.23	0.36	0.15

Table 1: Root mean square error of the inpainting results in some sequences from Sintel dataset [8] using different methods.

We only consider the operator $\mathcal{L}$ for the rest of the experiments. In order to evaluate the sensitivity of our video inpainting method to the optical flow accuracy, we present results with the same sequence but with different optical flows, computed using two different optical flow estimation methods. The seventh row displays the result computed with the optical flow estimated using the NLTV-CSAD energy functional. The eighth row uses the classical TV-L1 energy. In both cases the energy is minimized with a recently proposed strategy [30]. The results using the three different optical flows are very similar.

Experiments where an object is removed

Figs. 3 and 4 display experiments where we remove an object that is occluding another one. Fig. 3 is a sequence formed by the first 8 frames of the cave_2 sequence from Sintel, and Fig. 4 by 5 frames (12 to 16) of the temple_2 sequence. The object to be completed is shown in white in the second row. The inpainting mask, always depicted in gray in the second row, corresponds to the segmentation of the object we want to remove. Notice that this time, as a difference with the previous experiments, the inpainting mask is not static in time. In this case we also need to interpolate the optical flow in the inpainting domain and, for that, we use the method proposed in [29]. The completed optical flow is shown in the third row. The occlusion map is estimated from this completed flow using the criterion (6). Before applying the proposed method we inpaint, independently, the first and last frame with a 2D binary inpainting method (in particular, the perceptual-based 2D-inpainting method [27]). Then we apply the proposed video inpainting method that takes into account the estimated optical flow and occlusion map, and the completed shape is shown in the last row. Notice how, in Fig. 3, the hole corresponding to the spear is correctly filled and the dragon claw is partially recovered in frames 5 and 6 although it was completely occluded in the corresponding frames of the input sequence.

5. CONCLUSIONS

This work proposes a variational method for binary video inpainting. For that, we follow a threshold dynamics strategy where the dynamic shape analysis imposes spatial and temporal smoothness along the visible trajectory of the object by incorporating the convective derivative in a differential operator based on a generalized 3D gradient. Our proposal allows to keep track of the motion occlusions among the moving binary objects. We present some experimental results on the Sintel database containing complex object motion and occlusions. As future work we plan to study a variational model for the joint estimation of the shape and optical flow completion.

6. REFERENCES

[1] J.-F. Aujol, S. Ladjal, and S. Masnou. Exemplar-based inpainting from a variational point of view. *SIAM J. on Mathematical Analysis*, 42(3):1246–1285, 2010.

[2] C. Ballester, M. Bertalmío, V. Caselles, G. Sapiro, and J. Verdera. Filling-in by joint interpolation of vector fields and gray levels. *IEEE Trans. on IP*, 10(8):1200–1211, 2001.

[3] G. Barles and C. Georgelin. A simple proof of convergence for an approximation scheme for computing motions by mean curvature. *SIAM J. on Numerical Analysis*, 32:484–500, 1995.

[4] A. L. Bertozzi, S. Esedoglu, and A. Gillette. Inpainting of binary images using the cahn-hilliard equation. *IEEE Trans. on IP*, 16(1):285–291, 2007.

[5] P. Bhat, C. L. Zitnick, M. Cohen, and B. Curless. Gradientshop: A gradient-domain optimization framework for image and video filtering. *ACM Trans. on Graphics (TOG)*, 29(2):10:1–10:14, 2010.

[6] P. Bhat, C. L. Zitnick, N. Snavely, A. Agarwala, M. Agrawala, B. Curless, M. Cohen, and S. B. Kang. Using photographs to enhance videos of a static scene. In *Eurographics Symposium on Rendering (EGSR)*, pages 327–338, 2007.

[7] A. Bugeau, P. Gargallo, O. D'Hondt, A. Hervieu, N. Papadakis, and V. Caselles. Coherent background video inpainting through kalman smoothing along trajectories. In *Int. Workshop on Vision, Modeling, and Visualization Workshop*, pages 123–130, 2010.

[8] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black. A naturalistic open source movie for optical flow evaluation. In *ECCV*, pages 611–625, 2012.

[9] F. Cao, Y. Gousseau, S. Masnou, and P. Pérez. Geometrically guided exemplar-based inpainting. *SIAM J. on Imaging Sciences*, 4(4):1143–1179, 2011.

[10] T. Chan and J. H. Shen. Mathematical models for local nontexture inpaintings. *SIAM J. App. Math.*, 62(3):1019–43, 2001.

[11] G. Citti and A. Sarti. A cortical based model of perceptual completion in the roto-translation space. *J. of Math. Imaging and Vision*, 24(3):307–326, 2006.

[12] A. Criminisi, P. Pérez, and K. Toyama. Region filling and object removal by exemplar-based inpainting. *IEEE Trans. on IP*, 13(9):1200–1212, 2004.

[13] M. Ebdelli, O. Le Meur, and C. Guillemot. Video inpainting with short-term windows: application to object removal and error concealment. *IEEE Trans. on IP*, 24:3034–3047, 2015.

[14] S. Esedoglu, S. Ruuth, and R. Tsai. Threshold dynamics for shape reconstruction and disocclusion. In *IEEE Proc. of ICIP*, volume 2, pages 502–505, 2005.

[15] S. Esedoglu and Y.-H. R. Tsai. Threshold dynamics for the piecewise constant Mumford–Shah functional. *J. of Computational Physics*, 211(1):367–384, 2006.

[16] L. C. Evans. Convergence of an algorithm for mean curvature motion. *Indiana Univ. Mathematics J.*, 42:533–557, 1993.

[17] G. Facciolo, R. Sadek, A. Bugeau, and V. Caselles. Temporally consistent gradient domain video editing. In *Int. Workshop on Energy Min. Methods in Computer Vision and Pattern Recognition*, pages 59–73, 2011.

[18] M. Granados, J. Tompkin, K. Kim, O. Grau, J. Kautz, and C. Theobalt. How not to be seen object removal from videos of crowded scenes. In *Computer Graphics Forum*, volume 31, 2012.

[19] R. Grzhibovskis and A. Heintz. A convolution thresholding scheme for the willmore flow. *Interfaces and Free Boundaries*, 10(2):139–153, 2008.

[20] B. K. P. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[21] C. Kondermann, D. Kondermann, and C. Garbe. Postprocessing of optical flows via surface measures and motion inpainting. In *Joint Pattern Recognition Symposium*, pages 355–364. Springer Berlin Heidelberg, 2008.

[22] S. Masnou. Disocclusion: a variational approach using level lines. *IEEE Trans. on IP*, 11(2):68–76, 2002.

[23] S. Masnou and J.-M. Morel. Level lines based disocclusion. In *IEEE Proc. of ICIP*, 1998.

[24] O. E. G. W. T. X. Matsushita, Y. and H. Y. Shum. Full-frame video stabilization with motion inpainting. *IEEE Transactions on Pattern Analysis and Machine Intelligences*, 28(7):1150–1163, 2006.

[25] B. Merriman, J. Bence, and S. Osher. Diffusion generated motion by mean curvature. In *J.E. Taylor ed., Computational Crystal Growers Workshop*, pages 73–83. American Mathematical Society, 1992.

[26] A. Newson, A. Almansa, M. Fradet, Y. Gousseau, and P. Pérez. Video inpainting of complex scenes. *SIAM J. on Imaging Sciences*, 7:1993–2019, 2014.

[27] M. Oliver, G. Haro, M. Dimiccoli, B. Mazin, and C. Ballester. A computational model for amodal completion. *J. of Math. Imaging and Vision*, 56(3):1–24, 2016.

[28] K. A. Patwardhan, G. Sapiro, and M. Bertalmío. Video inpainting under constrained camera motion. *IEEE Trans. on IP*, 16(2):545–553, 2007.

[29] R. P.Palomares, G. Haro, and C. Ballester. A rotation-invariant regularization term for optical flow related problems. In *ACCV 2014*, pages 304–319. 2014.

[30] R. P.Palomares, E. Meinhardt-Llopis, C. Ballester, and G. Haro. Faldoi: A new minimization strategy for large displacement variational optical flow. *J. of Math. Imaging and Vision*, 2016.

[31] R. Sadek, G. Facciolo, P. Arias, and V. Caselles. A variational model for gradient-based video editing. *Int. J. of Computer Vision*, 103(1):127–162, 2013.

[32] P. Sand and S. Teller. Particle video: Long-range motion estimation using point trajectories. *Int. J. of Computer Vision*, 80(1):72–91, 2008.

[33] M. Strobel, J. Diebold, and D. Cremers. Flow and color inpainting for video completion. In *German Conference on Pattern Recognition*, pages 293–304, 2014.

[34] C. Vogel, S. Roth, and K. Schindler. An evaluation of data costs for optical flow. In *Pattern recognition*, pages 343–353. 2013.

[35] M. Werlberger, T. Pock, and H. Bischof. Motion estimation with non-local total variation regularization. In *IEEE Proc. of CVPR*, pages 2464–2471, 2010.

[36] Y. Wexler, E. Shechtman, and M. Irani. Space-time completion of video. *IEEE Trans. on PAMI*, 29(3):463–476, 2007.

[37] Z. Xu, Q. Zhang, Z. Cao, and C. Xiao. Video background completion using motion-guided pixels assignment optimization. *IEEE Trans. on CSVT*, 2015.

[38] C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime tv-l 1 optical flow. In *Pattern Recognition*, pages 214–223. 2007.