Spatio-Temporal Binary Video Inpainting via Threshold Dynamics

M. Oliver (maria.oliverp@upf.edu), R. P. Palomares, C. Ballester, G. Haro
Universitat Pompeu Fabra (Barcelona)

1 Introduction

- Variational method for the completion of moving shapes through binary video inpainting that works by smoothly recovering the objects into an inpainting hole.
- The model takes into account the optical flow and motion occlusions.
- The algorithm is based on threshold dynamics.

2 The Model

Let $u_0(\mathbf{x},t)$ be a binary video sequence defined on $\mathcal{V}\setminus\mathcal{M}$, where $\mathcal{V}=\{(\mathbf{x},t):\mathbf{x}=(x,y)\in\Omega,\ t\in\mathbb{R}\}$ is the video domain and $\mathcal{M}\subset\mathcal{V}$ denotes the inpainting hole with missing information. 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 the operator $\mathcal{L}(u)$ is

$$\mathcal{L}(u)=(u_x,u_y,\gamma\chi\partial_{\mathbf{v}}u)\,,\qquad \gamma>0. \tag{2}$$

If we make $\gamma\to\infty$, problem (1) is equivalent to

$$\min_{u:\mathcal{V}\to\{0,1\}}\int_{\mathcal{M}}\|\tilde{\mathcal{L}}(u)\|^2,\quad \text{s.t. } u=u_0 \text{ in } \mathcal{V}\setminus\mathcal{M}, \tag{3}$$

where,

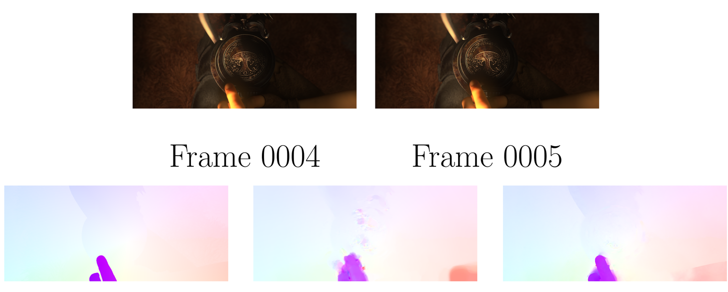
$$\tilde{\mathcal{L}}(u)=\partial_{\mathbf{v}}u. \tag{4}$$

- The first and last frames are inpainted using the inpainting model proposed in [3].
- We want the functional to only impose temporal regularity along the pixel trajectories that are not occluded. χ indicates the occlusion areas.
- The convective derivative is defined as

$$\partial_{\mathbf{v}}u(\mathbf{x},t)=\nabla u(\mathbf{x},t)\cdot\mathbf{v}(\mathbf{x},t)+\frac{\partial u}{\partial t}(\mathbf{x},t). \tag{5}$$

Optical Flow Estimation

1- Equation (5) involves the Optical Flow. We tried our model with different OF estimations.

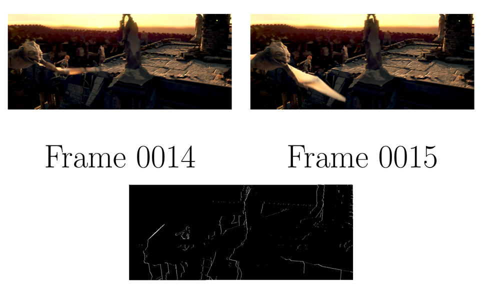


2- We also need to inpaint the optical flow in the inpainted regions.

Occlusion Estimation

We propose to use the method proposed in [6]:

$$\chi(\mathbf{x},t)=\begin{cases}1 & \text{if } \operatorname{div}(\mathbf{v})\geq -0.5 \text{ (visible at t+1)}\\ 0 & \text{else (not visible at t+1)}\end{cases} \tag{6}$$



3 Algorithm: Threshold Dynamics

We consider the equivalent problem:

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

where $\varepsilon>0$ and $W(u)=u^2\left(1-u^2\right)$.
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 $\chi\partial_{\mathbf{v}}$.
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 $\mathcal{T}=\left\{\mathbf{x}:\ \bar{u}(\tau)(\mathbf{x})\geq\frac{1}{2}\right\}$
3. **Fidelity step**. $\mathcal{T}^{n+1}=(\mathcal{T}\cap\mathcal{M})\cup\left(\mathcal{T}^0\cap(\mathcal{V}\setminus\mathcal{M})\right)$. We impose that the binary video coincides with the original video outside the inpainting domain.

4 Results

Experiment where a damaged object is recovered

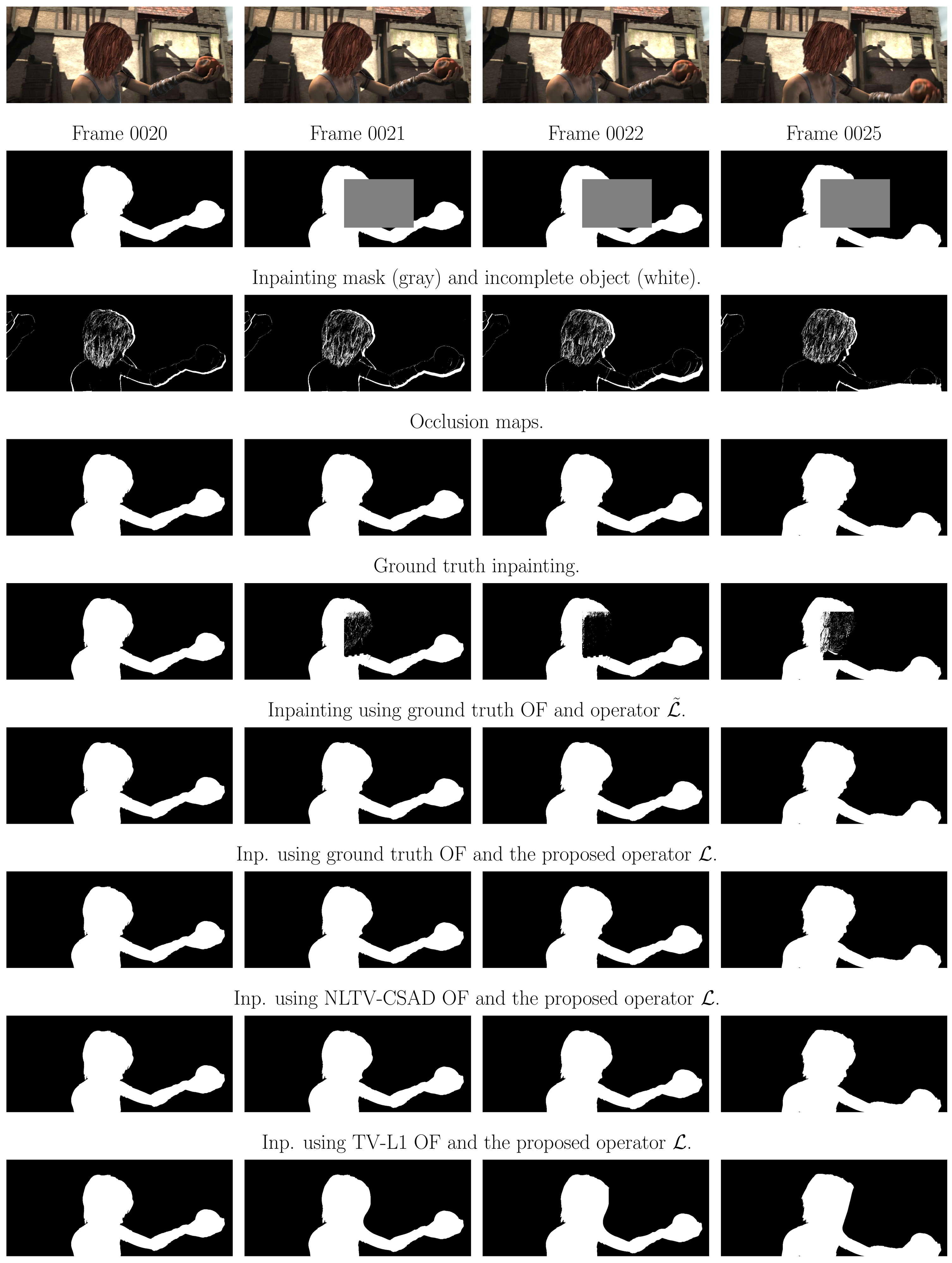


Fig. 4: Experiment with alley_1 sequence from Sintel [1]: Inpainting results with different methods and optical flow estimations.

Experiment where an object is removed

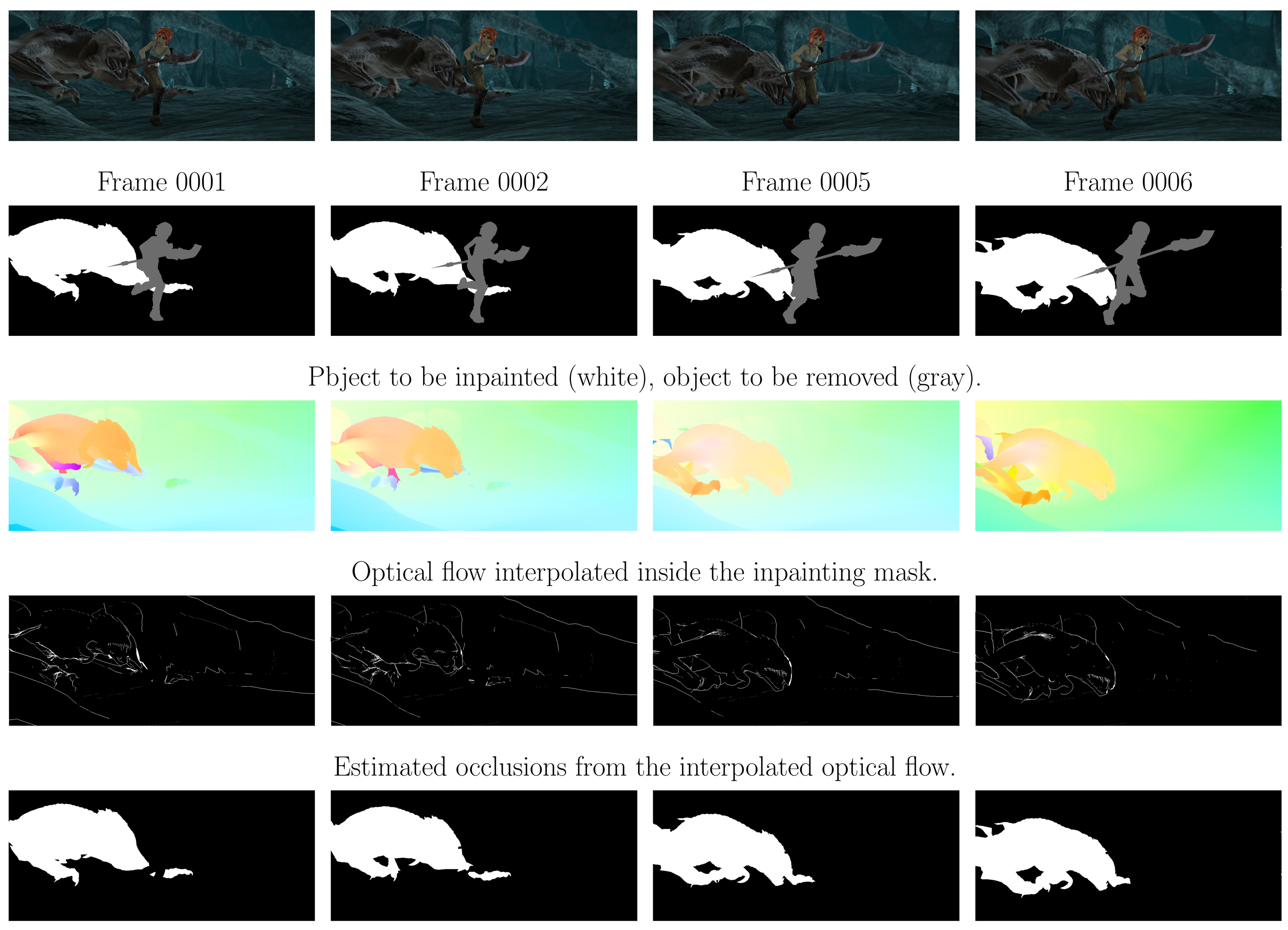


Fig. 5: Removal of an object in a video sequence (cave2).

5 Numerical Results

Root mean square error of the inpainting results in some sequences from Sintel dataset [1].

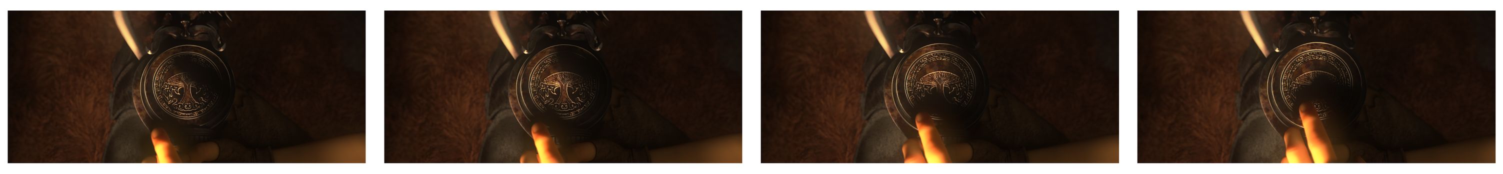
	MBO [2]	$\tilde{\mathcal{L}}$	$\mathcal{L}$
alley_1	0.18	0.55	**0.06**
ambush_4	0.46	0.54	**0.26**
market_5	0.34	0.23	**0.07**
shaman_3 (seq.1)	0.25	0.10	**0.05**
shaman_3 (seq.2)	0.63	0.63	**0.48**
temple_3	0.23	0.36	**0.15**

6 References

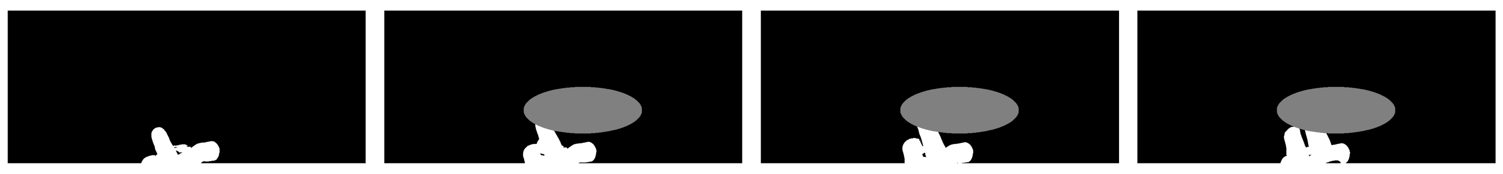
[1] D.J.Butler, J.Wulff, G.B.Stanley,M.J.Black (2012)
[2] B.Merriman,J.Bence,S.Osher (1992)
[3] M.Oliver, G.Haro, M.Dimiccoli, B.Mazin, C.Ballester (2016)
[4] R.P.Palomares, G.Haro, C.Ballester (2014)
[5] R.P.Palomares, E.Meinhardt-Llopis, C.Ballester, G.Haro (2016)
[6] P.Sand, S.Teller (2008)
[7] C.Zach,T.Pock,H.Bischof (2007)

Spatio-Temporal Binary Video Inpainting via Threshold Dynamics (continuation: More Results)

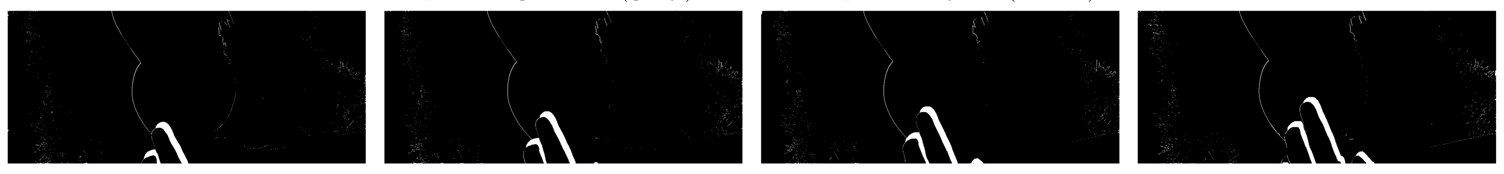
Experiments where a damaged object is recovered



Frame 0020 **Frame 0021** **Frame 0022** **Frame 0025**



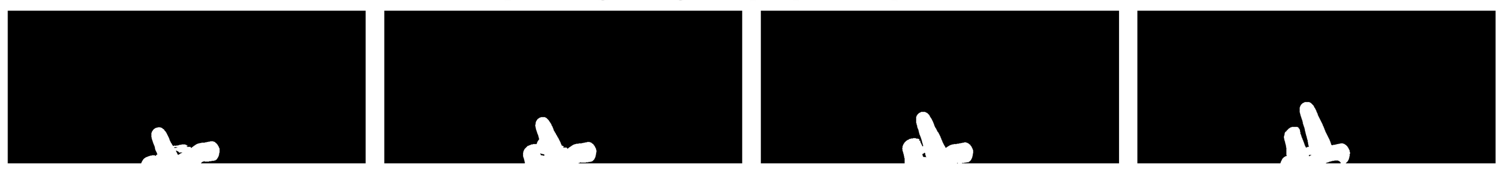
Inpainting mask (gray) and incomplete object (white).



Occlusion maps.

Groundtruth Inpainting results.

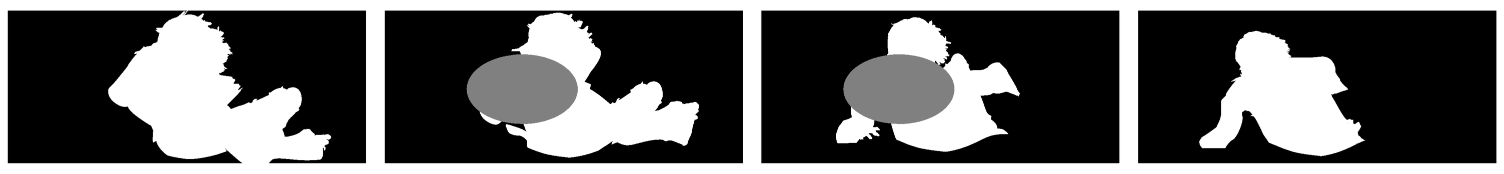
Inpainting using the proposed operator $\tilde{\mathcal{L}}$.



Inpainting using the proposed operator $\mathcal{L}$.

Inpainting results using 3D-MBO [2].

Frame 0015 **Frame 0016** **Frame 0018** **Frame 0020**



Inpainting mask (gray) and incomplete object (white).

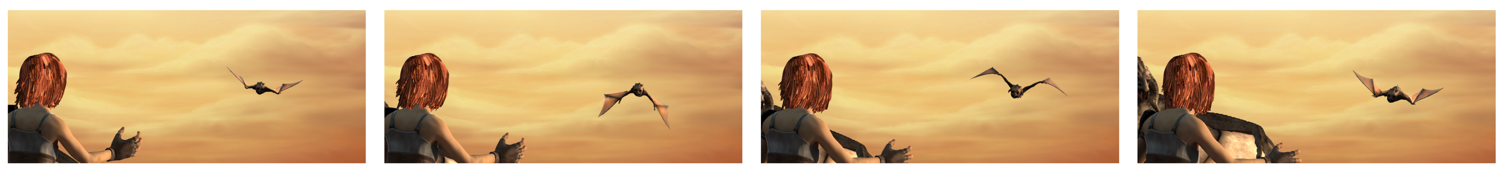
Occlusion maps.

Groundtruth Inpainting results.

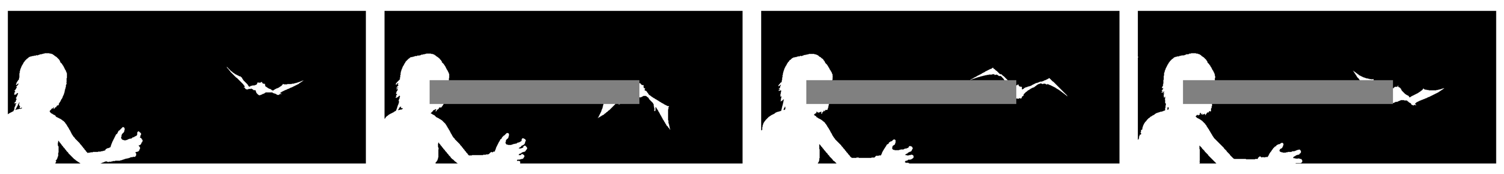
Inpainting using the proposed operator $\tilde{\mathcal{L}}$.

Inpainting using the proposed operator $\mathcal{L}$.

Inpainting results using 3D-MBO [2].



Frame 0010 **Frame 0012** **Frame 0014** **Frame 0016**



Inpainting mask (gray) and incomplete object (white).

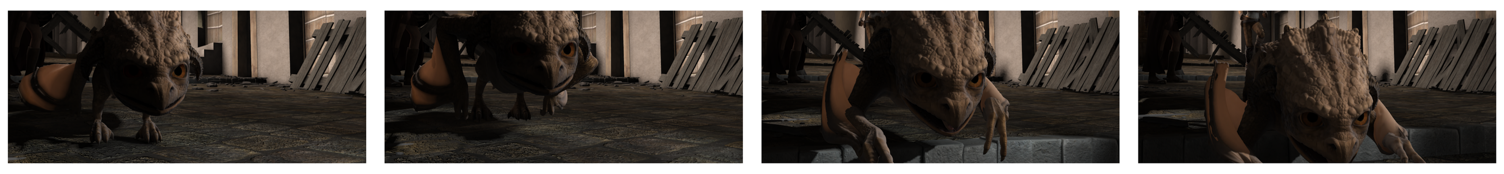
Occlusion maps.

Groundtruth Inpainting results

Inpainting using the proposed operator $\tilde{\mathcal{L}}$.

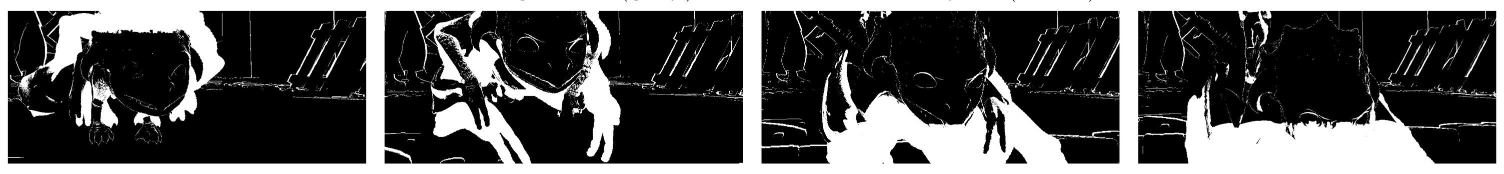
Inpainting using the proposed operator $\mathcal{L}$.

Inpainting results using 3D-MBO [2].



Frame 0010 **Frame 0012** **Frame 0014** **Frame 0016**

Inpainting mask (gray) and incomplete object (white).



Occlusion maps.

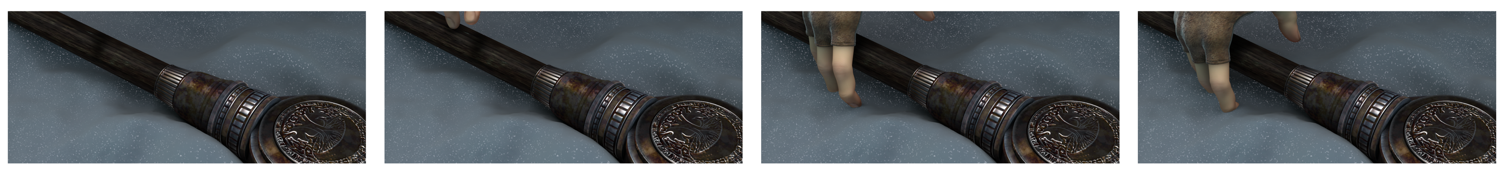
Groundtruth inpainting results.

Inpainting results using the proposed operator $\tilde{\mathcal{L}}$.

Inpainting results using the proposed operator $\mathcal{L}$.

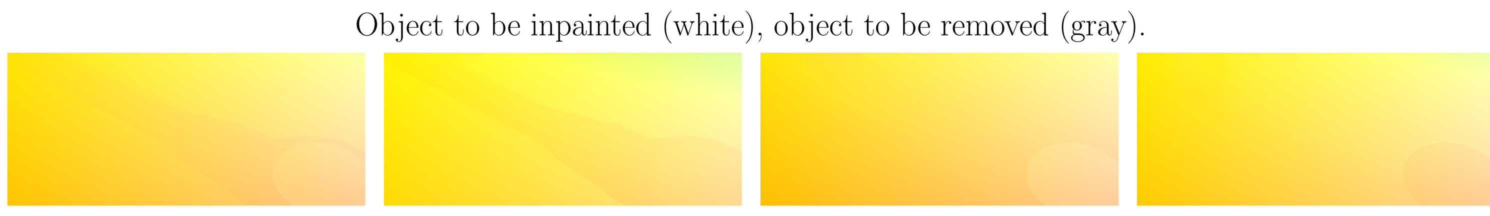
Inpainting results using 3D-MBO [2].

Experiments where an object is removed

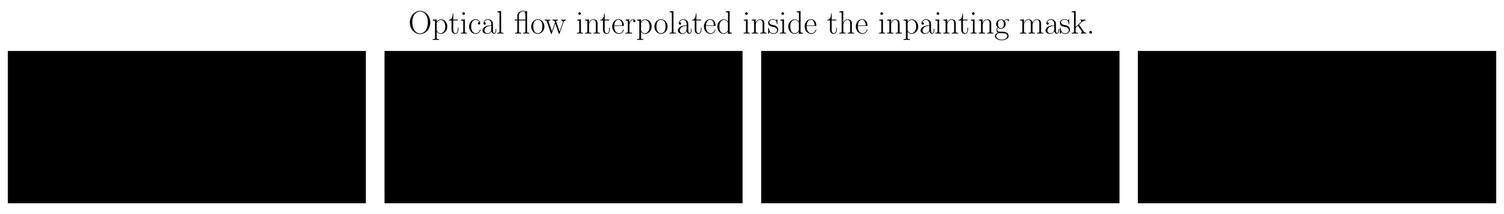


Frame 0004 **Frame 0005** **Frame 0009** **Frame 0010**

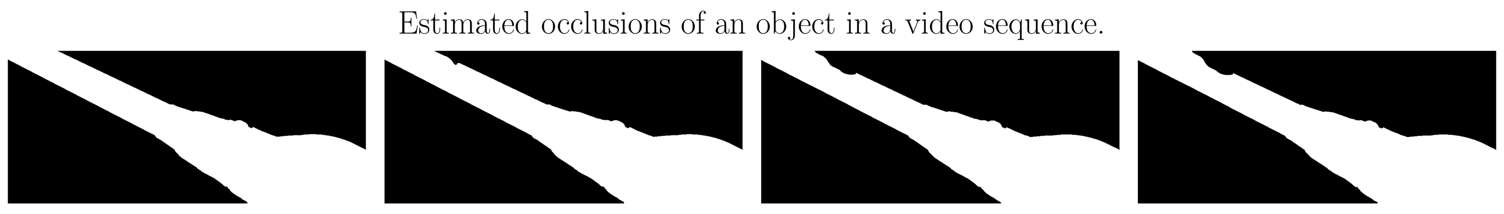
Object to be inpainted (white), object to be removed (gray).



Optical flow interpolated inside the inpainting mask.



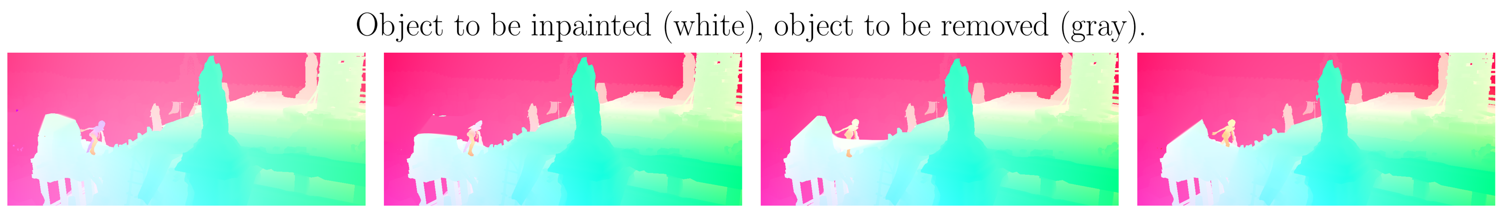
Estimated occlusions of an object in a video sequence.



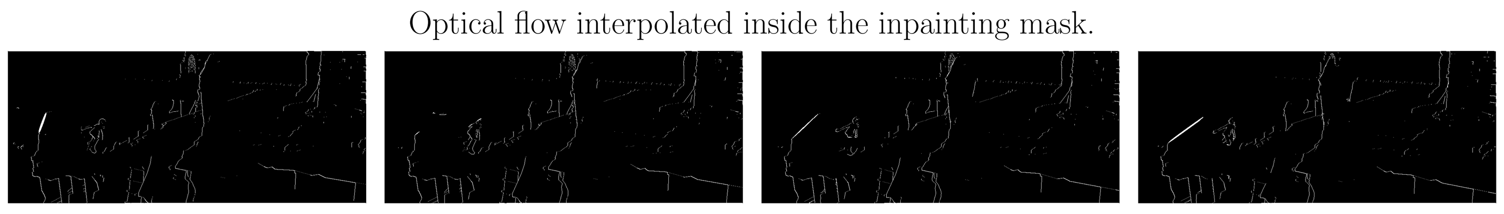
White object inpainted.

Frame 0012 **Frame 0013** **Frame 0014** **Frame 0015**

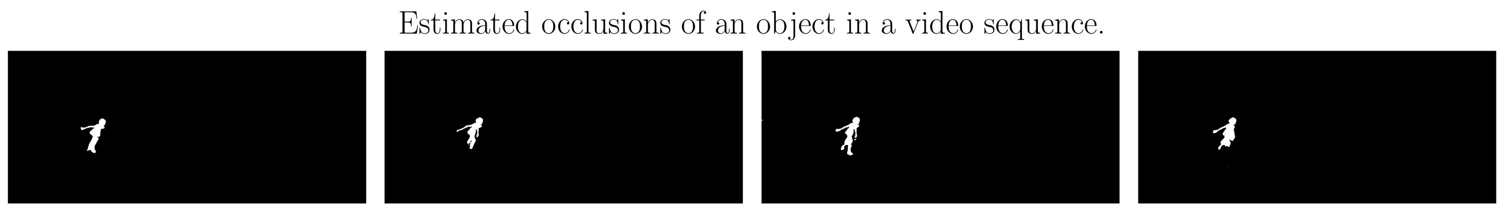
Object to be inpainted (white), object to be removed (gray).



Optical flow interpolated inside the inpainting mask.



Estimated occlusions of an object in a video sequence.



White object inpainted.