Temporally consistent gradient based video editing

Gabriele Facciolo

Technicolor - Rennes - Oct 2013

Collaborators

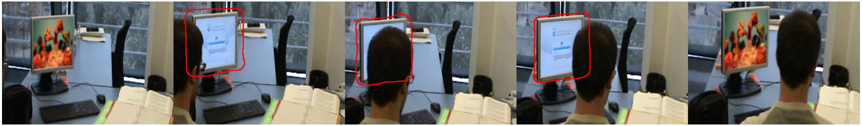
- G.F, R. Sadek, A. Bugeau, and V. Caselles. "Temporally Consistent Gradient Domain Video Editing". *EMMCVPR 2011, Saint-Petersburg.*
- R. Sadek, G.F, P. Arias, and V. Caselles. "A Variational Model for Gradient-Based Video Editing". *IJCV 2013.*

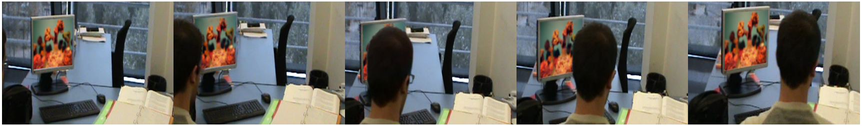
Motivation

sources/destinations cloning seamless cloning

[Pèrez et al. ’03]

Motivation





Outline

Model

Discretization

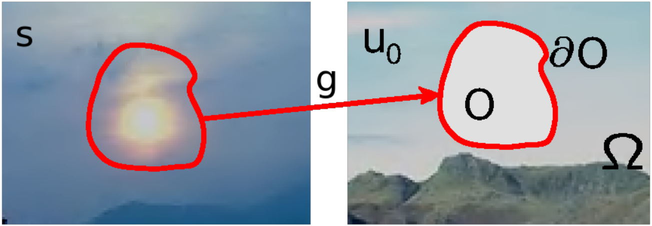
Results

Perspectives and Conclusion

Poisson edit

$$\min_u \int_{O \subset \Omega} \|\nabla u - \mathbf{g}\|^2 \mathrm{d}x \qquad \text{with} \qquad \underbrace{u|_{\partial O} = u_0}_{\text{Dirichlet boundary cond.}}$$

- $O \subset \Omega \subset \mathbb{R}^2$: edit domain
- $u : O \to R$: solution image
- $u_0 : \Omega \to \mathbb{R}$: target image
- $\mathbf{g} : O \to \mathbb{R}^2$: guidance field (i.e. ∇s)



[Pérez et at. 03]

Poisson edit

$$\min_u \int_{O\subset\Omega} \|\nabla u - \mathbf{g}\|^2 \mathrm{d}x \qquad \text{with} \qquad \underbrace{u|_{\partial O} = u_0}_{\text{Dirichlet boundary cond.}}$$

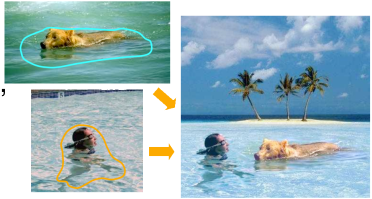
- $O \subset \Omega \subset \mathbb{R}^2$: edit domain
- $u : O \to R$: solution image
- $u_0 : \Omega \to \mathbb{R}$: target image
- $\mathbf{g} : O \to \mathbb{R}^2$: guidance field (i.e. ∇s)

[Pérez et at. 03]

From Image to Video Edit

Gradient domain image edit [Pèrez et. al., Georgiev, Hagenburg et. al.]...

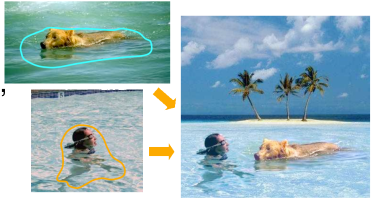
- manipulate image gradients
- applications: seamless compositing, tone mapping, shadow removal, matting, inpainting...



From Image to Video Edit

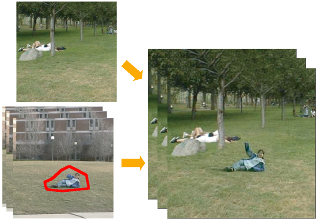
Gradient domain image edit [Pèrez et. al., Georgiev, Hagenburg et. al.]...

- manipulate image gradients
- applications: seamless compositing, tone mapping, shadow removal, matting, inpainting...



Extension to video [Wang et. al., Bhat et. al.]...

- applications: object insertion /removal, nonphotorealistic rendering, flickering removal...
- key: temporal consistency



Unwrap Mosaics: A new representation for video editing*

Alex Rav-Acha
Weizmann Institute of Science

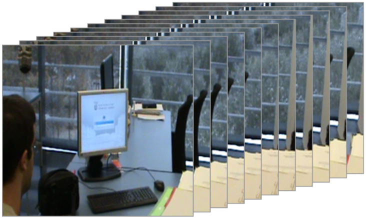
Pushmeet Kohli
Microsoft Research

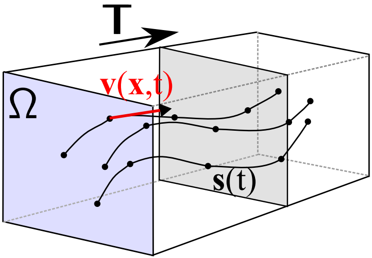
Carsten Rother
Microsoft Research

Andrew Fitzgibbon
Microsoft Research

Modeling Temporal Consistency

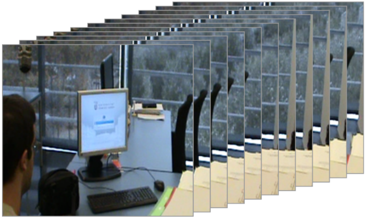
- Video as a volume / stack of frames $u : (\Omega \times \mathbb{T}) \to \mathbb{R}$
- Points **should "look uniform"** along their trajectories $\mathbf{s}(t)$

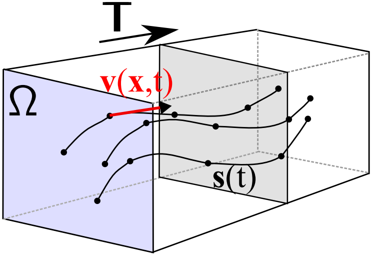




Modeling Temporal Consistency

- Video as a volume / stack of frames $u : (\Omega \times \mathbb{T}) \to \mathbb{R}$
- Points **should "look uniform"** along their trajectories $\mathbf{s}(t)$





- The motion field $\mathbf{v}(\mathbf{x}, t) : (\Omega \times \mathbb{T}) \to \mathbb{R}^2$ characterizes all the trajectories

Brightness Constancy Assumption

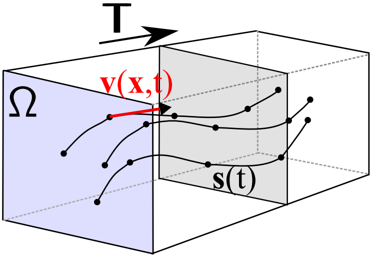
Points should have constant color along their trajectories $\mathbf{s}(t)$

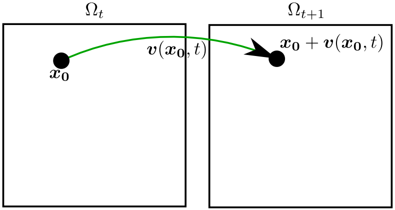
$$\frac{\mathrm{d}}{\mathrm{d}t} u(\mathbf{s}(t), t) = 0$$

Applying chain's rule

$$\partial_{\mathbf{v}} u(\mathbf{x}, t) := \mathbf{v}(\mathbf{x}, t) \cdot \nabla u(\mathbf{x}, t) + \partial_t u(\mathbf{x}, t) = 0$$

where $\partial_{\mathbf{v}} u$ is the convective derivative





Brightness Constancy Assumption

Example of brightness NON constancy

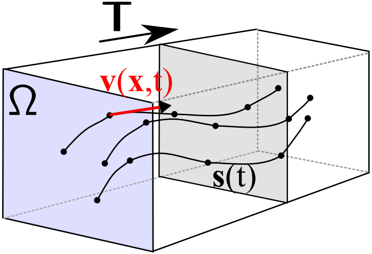
Global Brightness Change Assumption

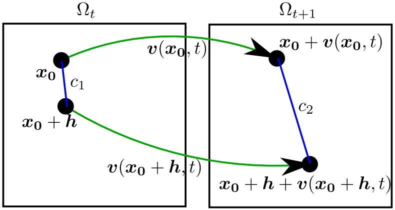
Points “look uniform” along their trajectories even in presence of a global brightness change

$$\partial_{\mathbf{v}} u(\mathbf{x}, t) = g(t)$$

Taking the spatial gradient on both sides

$$\boldsymbol{\nabla} \partial_{\mathbf{v}} u(\mathbf{x}, t) = 0$$





Modeling Motion

Feature Tracking + Parametric Model

1. Tracking: “Lagrangian” view of the motion

Modeling Motion

Optical Flow

1. Tracking: “Lagrangian” view of the motion
2. Optical flow: “Eulerian” view of the motion

Modeling Motion

Optical Flow

1. Tracking: "Lagrangian" view of the motion
2. Optical flow: "Eulerian" view of the motion $\Leftarrow$

Optical Flow vs. True Motion

True Motion (motion field)

object motion projected on the image plane

Optical Flow

estimation of the apparent motion (velocity) of objects within an image sequence

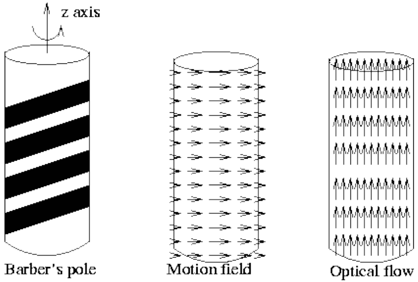
Optical Flow vs. True Motion

True Motion (motion field)

object motion projected on the image plane

Optical Flow

estimation of the apparent motion (velocity) of objects within an image sequence



http://homepages.inf.ed.ac.uk/rbf/CVonline

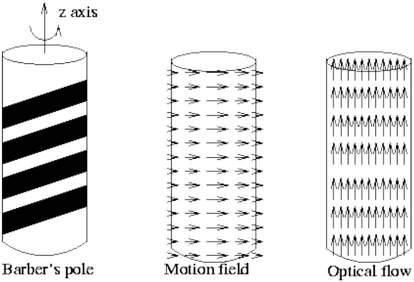
Optical Flow vs. True Motion

True Motion (motion field)

object motion projected on the image plane

Optical Flow

estimation of the apparent motion (velocity) of objects within an image sequence



http://homepages.inf.ed.ac.uk/rbf/CVonline

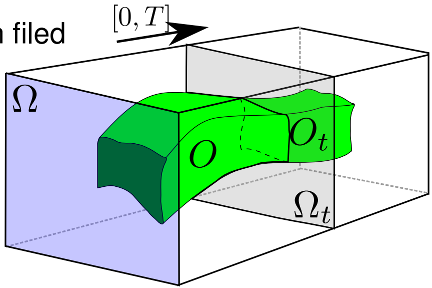
We use the optical flow as an approximation of the true motion

Gradient Based Video Edit (GBVE)

$$\min_{u} \int_0^T \int_{O_t} \|\nabla \partial_{\mathbf{v}} u(\mathbf{x}, t)\|^2 \, \mathrm{d}\mathbf{x} \, \mathrm{d}t$$

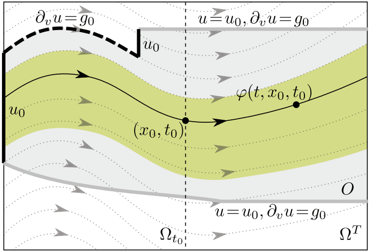
s.t. $u|_{\partial O} = u_0$

- $O \subset (\Omega \times \mathbb{T}) \subset \mathbb{R}^3$: edit domain
- $u_0 : (\Omega \times \mathbb{T}) \to \mathbb{R}$: target video
- $\mathbf{v} : (\Omega \times \mathbb{T}) \to \mathbb{R}^2$: motion filed
- $u : O \to \mathbb{R}$: solution

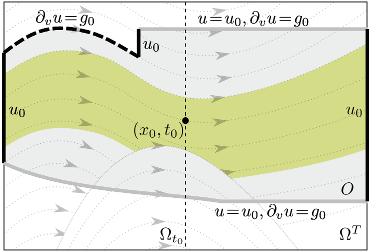


Gradient Based Video Edit (GBVE)

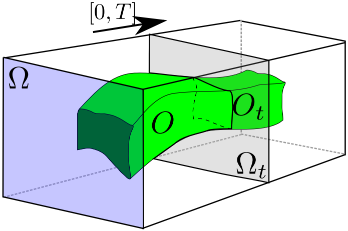
Boundary conditions



One-lid problem.



Two-lid problem.



Discretization

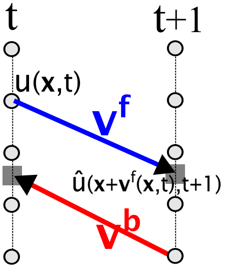
$$\min_{u} \sum_{(\boldsymbol{x},t)\in\widetilde{O}} \|\boldsymbol{\kappa}(\boldsymbol{x},t)\nabla\partial_{\boldsymbol{v}}u(\boldsymbol{x},t)\|^2 \quad + \quad BC$$

- Equivalent to a sparse linear least squares: $\min_u \|Au - b\|^2$
- Forward and backward flows $\boldsymbol{v}^f(\boldsymbol{x},t)$ and $\boldsymbol{v}^b(\boldsymbol{x},t)$, yield two convective derivative discretizations

$$\partial^f_{\boldsymbol{v}}u(\boldsymbol{x},t) := \hat{u}(\boldsymbol{x}+\boldsymbol{v}^f(\boldsymbol{x},t), t+1) - u(x,t)$$

$$\partial^b_{\boldsymbol{v}}u(\boldsymbol{x},t) := u(\boldsymbol{x},t) - \hat{u}(\boldsymbol{x}+\boldsymbol{v}^b(\boldsymbol{x},t), t-1)$$

where $\hat{u}(\boldsymbol{x},t)$ is the interpolation of u



Discretization

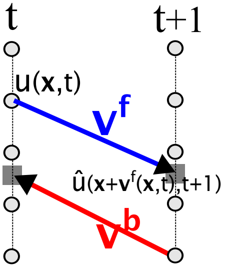
$$\min_{u} \sum_{(\boldsymbol{x},t)\in\widetilde{O}} \|\kappa(\boldsymbol{x},t)\nabla\partial_{\boldsymbol{v}} u(\boldsymbol{x},t)\|^2 \quad + \quad BC$$

- Equivalent to a sparse linear least squares: $\min_u \|Au - b\|^2$
- Forward and backward flows $\boldsymbol{v}^f(\boldsymbol{x},t)$ and $\boldsymbol{v}^b(\boldsymbol{x},t)$, yield two convective derivative discretizations

$$\partial^f_{\boldsymbol{v}} u(\boldsymbol{x},t) := \hat{u}(\boldsymbol{x} + \boldsymbol{v}^f(\boldsymbol{x},t), t+1) - u(x,t)$$

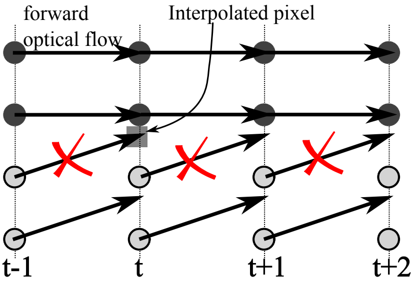
$$\partial^b_{\boldsymbol{v}} u(\boldsymbol{x},t) := u(\boldsymbol{x},t) - \hat{u}(\boldsymbol{x} + \boldsymbol{v}^b(\boldsymbol{x},t), t-1)$$

  where $\hat{u}(\boldsymbol{x},t)$ is the interpolation of u



- $\kappa(\boldsymbol{x},t)$: occlusion tensor

Occlusions Handling



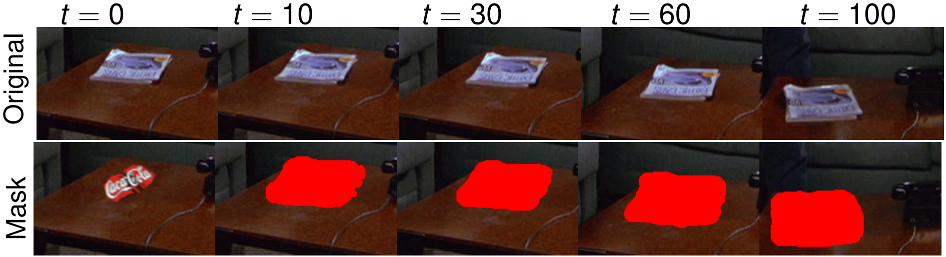
Ω_{t-1}

Ω_t

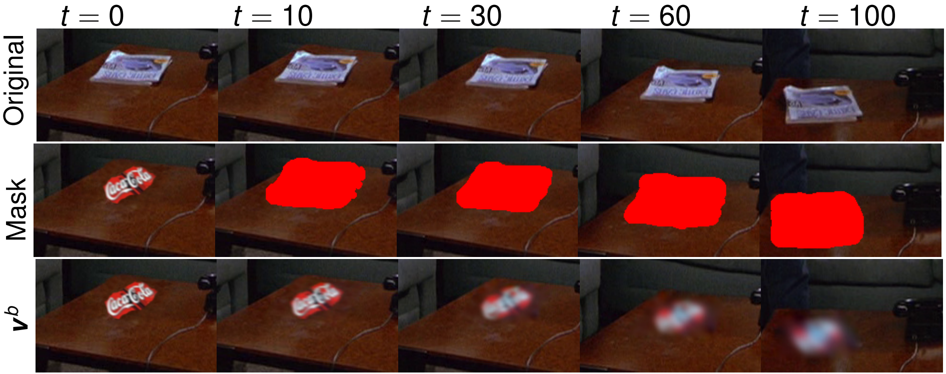
occlusions on Ω_{t-1}

$\kappa(\boldsymbol{x}, t)$ removes flow vectors corresponding to occlusions in u_0

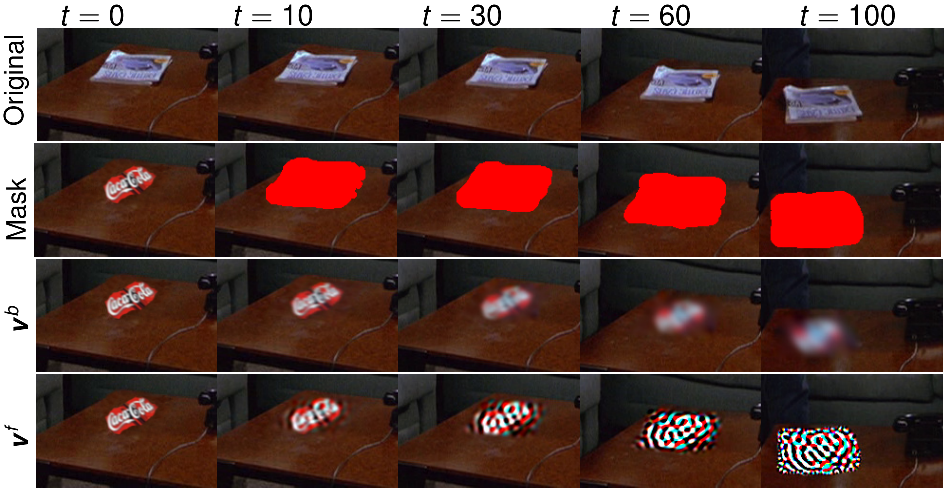
First GBVE experiment



First GBVE experiment



First GBVE experiment

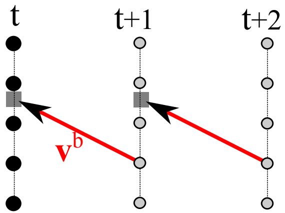


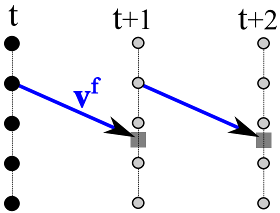
Analysis of discretized convective derivatives

- Constant translation motion, without brightness changes
- Suppose zero-energy solutions, so: $\partial_{\boldsymbol{v}}^{b/f} u(\boldsymbol{x}, t) = 0 \quad \forall(\boldsymbol{x}, t)$
- Denote interpolations M^b and M^f, so that

$$u(\cdot, t+1) = M^b\, u(\cdot, t) \qquad \Leftarrow \text{(blur)}$$

$$u(\cdot, t+1) = (M^f)^{-1}\, u(\cdot, t) \qquad \Leftarrow \text{(inverse smoothing)}$$





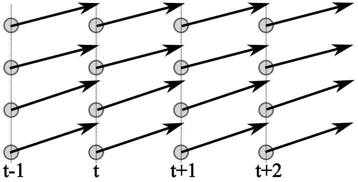
Key Obs.: $M^b$and $(M^f)^{-1}$ have somewhat opposed effects

DSCD: Deblurring Scheme for the Convective Derivative

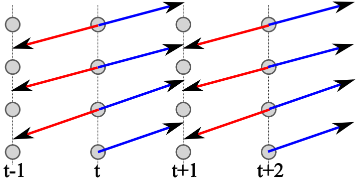
The idea is to **alternate** between

- $\mathbf{v}^f$-scheme at even frames (sharpens) and
- $\mathbf{v}^b$-scheme at odd frames (blurs)

to moderate each other's effects

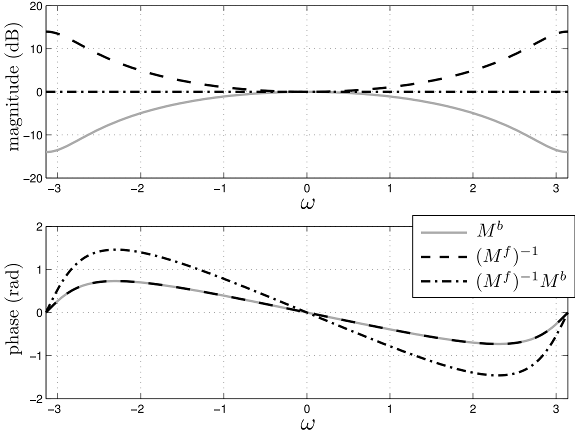


$\mathbf{v}^f$-scheme



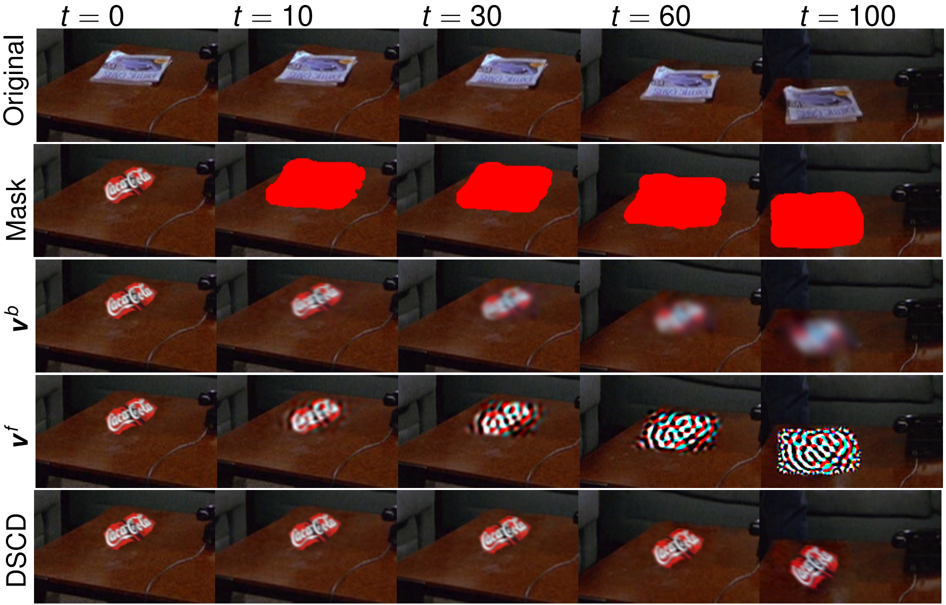
DSCD

DSCD

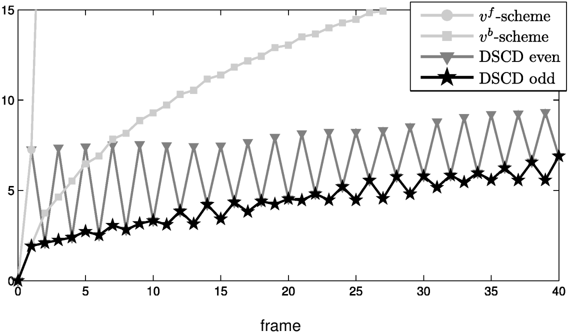


Analysis of a zero-energy solution for a one-lid problem, with a purely translational motion. Frequency responses of the M^b and $(M^f)^{-1}$ filters, and their composition $(M^f)^{-1}M^b$.

GBVE with DSCD

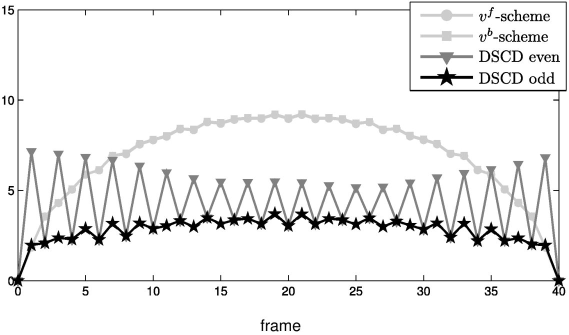


DSCD one-lid



RMSE w.r.t. the ground truth for a synthetic problem with a constant translation of $\boldsymbol{v_0} = [-0.425, 0]$ px/frame.

DSCD two-lid



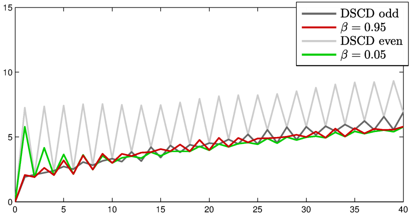
RMSE w.r.t. the ground truth for a synthetic problem with a constant translation of $\boldsymbol{v_0} = [-0.425, 0]$ px/frame.

DSCD - gory details

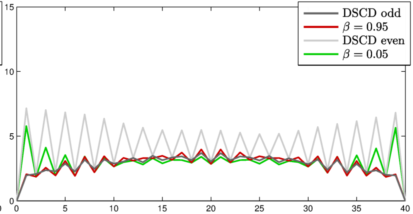
In practice, optimize a convex combination of the even and odd models

$$E_\beta = \beta E_{\boldsymbol{\kappa}}^{\text{odd}} + (1-\beta) E_{\boldsymbol{\kappa}}^{\text{even}},$$

with $\beta = 0.05$



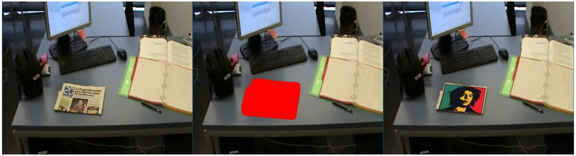
[One-lid case.]



[Two-lids case]

Results 1

Illumination chenages

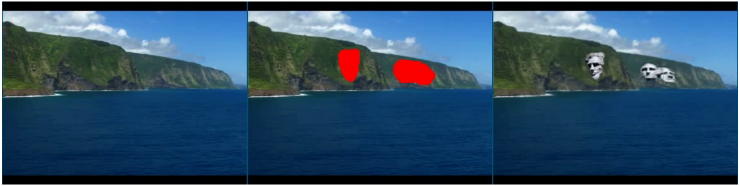


brightness constancy assumption

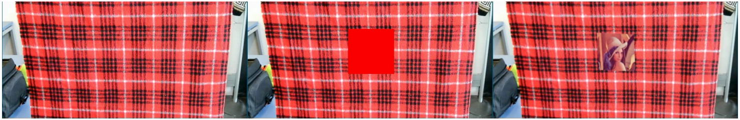
global brightness change assumption

Results 2

Deformations & Zoom



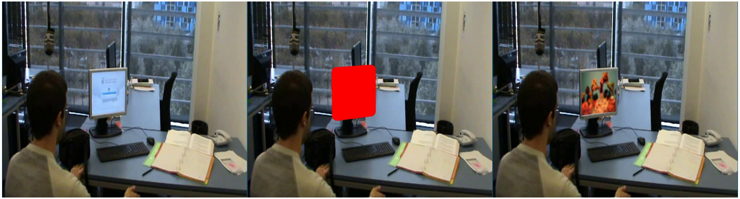
cliff



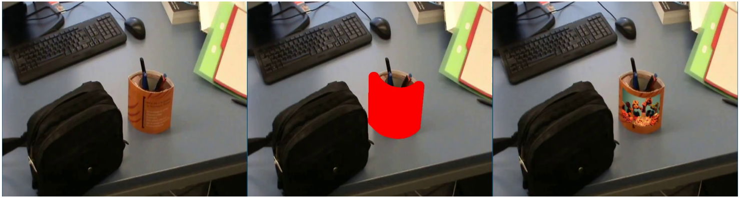
cloth

Results 3

Occlusions



screen



pen-holder

Discussion

- The GBVE algorithm relies extensively on the optical flow, unfortunately getting an artifact-free optical flow is hard.
- Sometimes tracking is easier (i.e. with markers).
  The Lagrangian modeling is more adequate when tracking can be done and motion is rigid.
- GBVE performs well specially with non-rigid objects, if the flow is available. While extracting trajectories from the field introduces drift.

Discussion

<table>
<tr><th>Eulerian modeling</th><th>Lagrangian modeling</th></tr>
<tr><td>Pros:
no need for parametric model
sometimes flow is available but tracking is not feasible</td><td>Pros:
precise for parametric motions
no long term degradation</td></tr>
<tr><td>Cons:
long term degradation
dependent on the quality of flow</td><td>Cons:
requires markers to track
drift if extracted from flow
artifacts if wrong parametric model is used</td></tr>
</table>

Perspectives

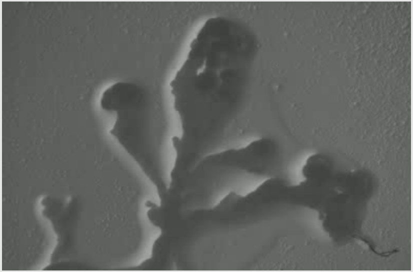
Use the Global Brightness Change Assumption

[The Cure, Charlie Chaplin, 1917]

Perspectives

Use the Global Brightness Change Assumption

to estimate motion in biological sequences



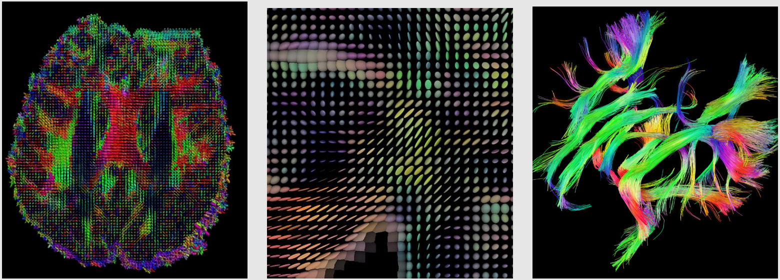
[Marc Durand (MSC Lab., Paris Diderot University),

http://perso.telecom-paristech.fr/∼delon/Demos/Flicker_stabilization/]

Perspectives

DSCD for simulation?

Diffusion tensor imaging (diffusion MRI) produces an Eulerian description of a field



[http://www.humanconnectomeproject.org/gallery/]

Conclusion

- Presented a functional for gradient domain video editing with temporal consistency
- Proposed Deblurring Scheme for Convective Derivative: to reduce artifacts introduced by the discretization of the temporal consistency term
- Discussed limitations and possible uses

Conclusion

- Presented a functional for gradient domain video editing with temporal consistency
- Proposed Deblurring Scheme for Convective Derivative: to reduce artifacts introduced by the discretization of the temporal consistency term
- Discussed limitations and possible uses

Thank you.

More results and MATLAB implementation available @
`http://gpi.upf.edu/static/gbve12/`