# Depth-aware Video Magnification

**Julian F. P. Kooij**[1,2] and **Jan C. van Gemert**[1]   [1]*Delft University of Technology*, [2]*Leiden University Medical Center*, *The Netherlands*

TU Delft   Leids Universitair Medisch Centrum

## ① Contributions

- Use depth cue to magnify motion of occluded regions
- Depth-Aware Steerable Pyramids
- Generalize the Fast Bilateral Filter to Non-Gaussian bilateral filters
- Application on and RGB+D dataset for tremors measurement

## ② Main motivation

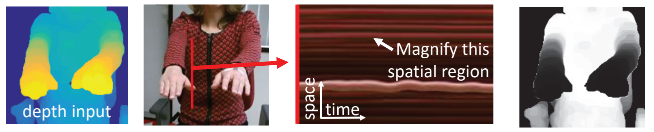
Medical application of full body tremor assessment

- Real-world hospital setting (e.g. Parkinson patients)
- Need to discover and measure *small* motions in arms, body, head, *with minimum patient effort*
- Should be robust against viewpoint, self-occlusions, and presence of large motions

Other uses of our novel filter explored in Sup. Mat.

## ③ Example magnification task

Magnify small motions in body, but large movements in foreground

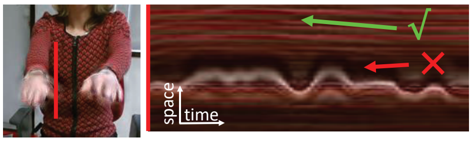



Input: **RGB+Depth** video from Kinect2. Subjects have **postural tremor**

Input **space-time slice** (see red line in RGB image)

Magnification matte defined on depth

## ④ Magnification comparison to state-of-the-art




**[1] Wadhwa et al.,SIGGRAPH'13**

Phase-based motion magnification:

- Per frame, build complex steerable pyramid
- Amplify temporal variations of complex pyramid coefficients
- Reconstruct video from amplified pyramids

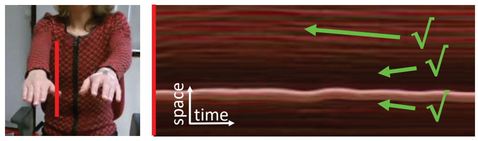
**Problem: magnifies small & large motions equally**


**[2] Elgharib et al., CVPR'15**

Magnify foreground with large motion:

- Stabilize large motions from camera
- Manually labeling of foreground to magnify
- Blend original and magnified video

**Problem: Large foreground movements still "leak" into magnified background region**




**[Ours] Depth-aware motion magnficiation**

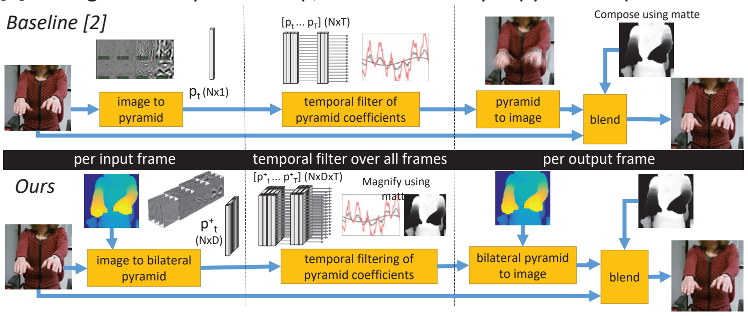
Construct per frame a "**bilateral pyramid**":

- Bilateral filter does depth-aware Gaussian blur
- But complex steerable filters are *non-Gaussian*
- **novel non-Gaussian bilateral filter**
  → ignores intensity at distant depth layers

No manual segmentation. **No foreground "leaking"**.

## ⑤ Processing pipeline comparison

[2] uses fg.mask only at last step; we use it directly in pyramid representation




## ⑥ Measuring motion task

Steerable Pyramids also used for motion measurement

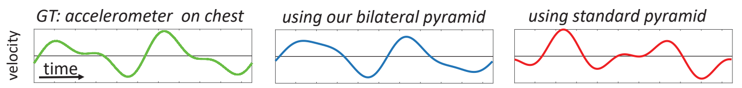
- "Leaking" into background affects measurement too
- Using our bilateral pyramid is therefore more robust

Task: measure the *chest area* behind the moving hand

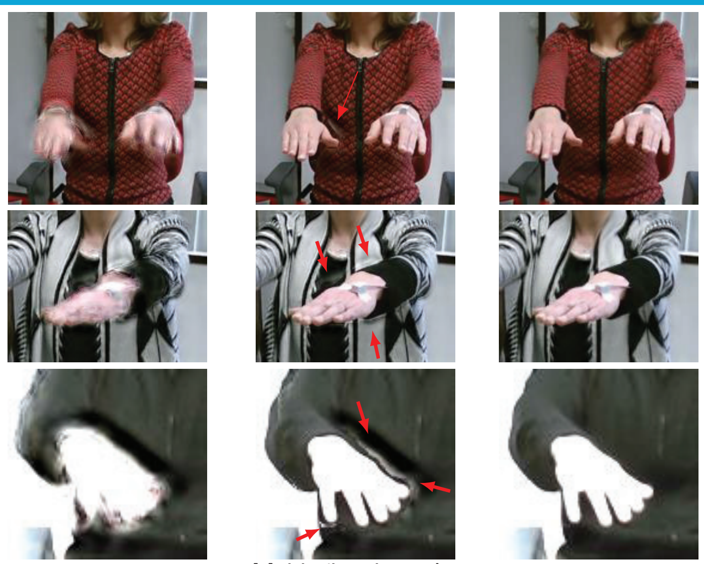
- Compare to accelerometer on chest and hand
- Standard pyramid: measures leaked motion of hand
- Bilateral pyramid: measures chest like accelerometer

measure **chest** area

standard pyramid motion `**leakage**'




## ⑦ More single frame magnification comparisons



**[1] Wadhwa et al., SIGGRAPH'13**
Magnifying small motions also distorts large fireground motions

**[2] Elgharib et al., CVPR'15**
Mask restores foreground, but artifacts in bg. remain

**Our approach**
Using depth-aware pyramid representation curbs artifacts

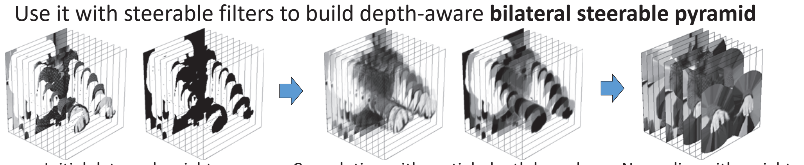
## ⑧ Building a bilateral steerable pyramid

Our novel non-Gaussian bilateral filter generalizes the *Fast Bilateral Filter [3]*

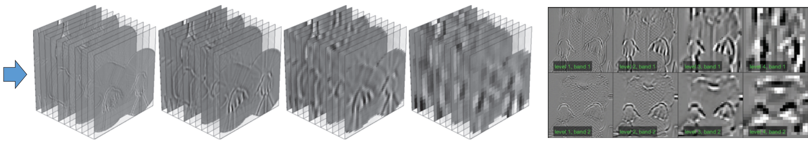
- Given input image *I(x)*, depth image *E(x)*, and let *x*, *y*, *z* be 2D image locations
- The **standard bilateral filter** outputs *O(x)*, using Gaussian kernel $G(d\,;\sigma)$
  - $O(x) = \frac{1}{W(x)} \sum_{y \in N(x)} w(|x-y|, E(x) - E(y))\, I(y)$
  - $w(d_s, d_E) = G(d_s; \sigma_s) \times G(d_E; \sigma_r)$
- Our **non-Gaussian bilateral filter** *Q(z)* for non-Gaussian kernels *F(d)*
  - $Q(z) = \sum_{y \in N(x)} F(|x-z|)\, O^+(x, E(z))$
  - $O^+(x,\xi) = \frac{1}{W^+(x,\xi)} \sum_{y \in N(z)} w(|x-y|, \xi - E(y))\, I(y)$

  Here $O^+(x,\xi)$ is a *volumatric* representation (2D image loc + 1D depth)
  → *Q* convolves *F* on all depth layers of $O^+$,which locally downweights the input at distant depth layers through a 3D kernel *w*

Use it with steerable filters to build depth-aware **bilateral steerable pyramid**



Initial data and weights   Convolution with spatial+depth kernel   Normalize with weights

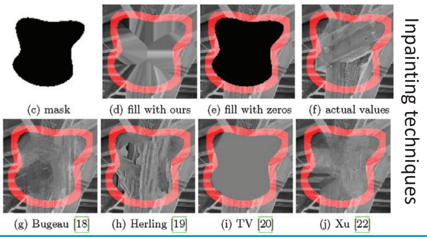


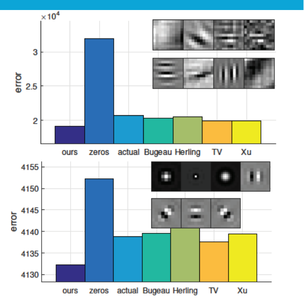
Apply *non-Gaussian steerable filters*, and rescale multiple times   Pyramid levels & orientations

## ⑨ Non-Gaussian bilateral experiments

Study non-Gauss filters on images + binary mask

- Ideally, filter **ignores** intensity within masked region
- Compare our method to using **inpainting** techniques
- Tested on steerable filters and ConvNet filters




**Result**: Inpainting yields large responses; Our method is fast, has minimum response

## ⑩ References, acknowledgements, and code

- **[1]** Wadhwa, N., Rubinstein, M., Durand, F., Freeman, W.T.: *Phase-based video motion processing*. Proceedings SIGGRAPH (2013), 32(4)
- **[2]** Elgharib, M.A., Hefeeda, M., Durand, F., Freeman, W.T.: *Video magnication in presence of large motions*. Proceedings CVPR (2015), p. 4119 - 4127
- **[3]** Paris, S., Durand, F.:*A fast approximation of the bilateral filter using a signal processing approach*. IJCV (2009) 81(1), p. 24 - 52

**Acknowledgements:** This work is part of the research programme Technology in Motion (TIM [628.004.001]), financed by the Netherlands Organisation for Scientific Research (NWO)

**Code: github.com/jkooij/depthaware-momag**   **Project page: tim.lumc.nl/**