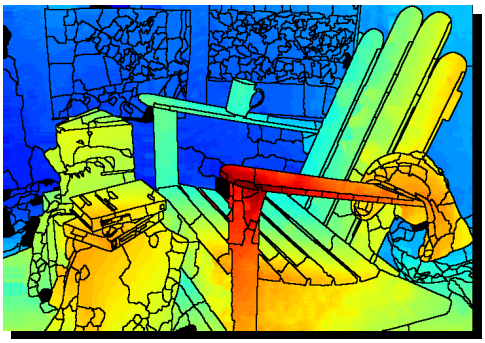
JOURNÉES DOCTORALES DES 22-24 JUIN 2016

MORPHOLOGICAL PROCESSING OF STEREOSCOPIC IMAGE SUPERIMPOSITIONS FOR DISPARITY MAP ESTIMATION



Jean-Charles Bricola, Michel Bilodeau & Serge Beucher

This work has been performed in the project PANORAMA, co-funded by grants from Belgium, Italy, France, the Netherlands, and the United Kingdom, and the ENIAC Joint Undertaking.

THE PROBLEM

Given two ***rectified*** stereo images...

AustraliaP / Left view

... compute a depth map from ***corresponding pixels***.

The Problem

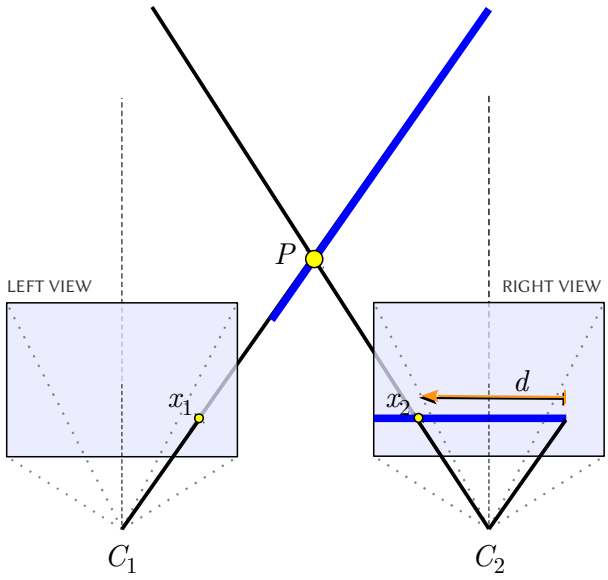
Given two ***rectified*** stereo images...

`AustraliaP` / Right view

... compute a depth map from ***corresponding pixels***.

Stereo Image Geometry

Relations between depth and disparity ***in the rectified case***:

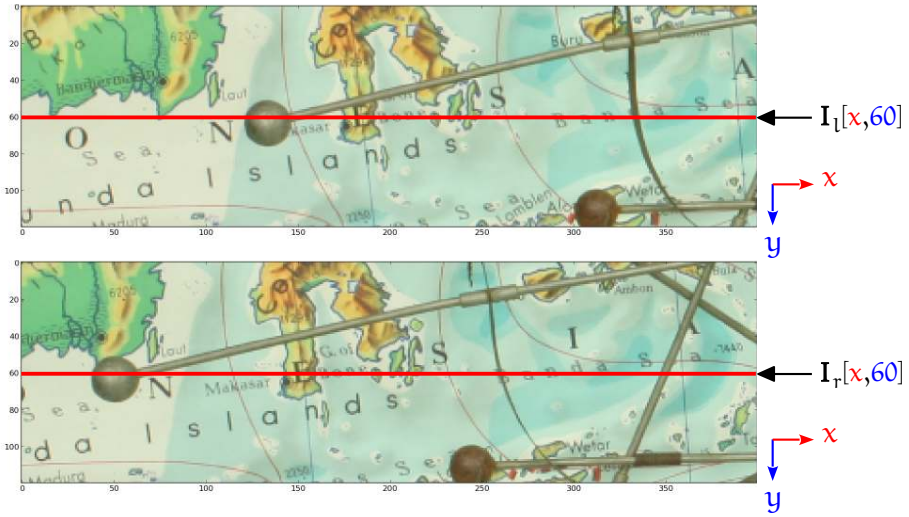


Correspondence Establishment & Superimpositions of Stereo Images

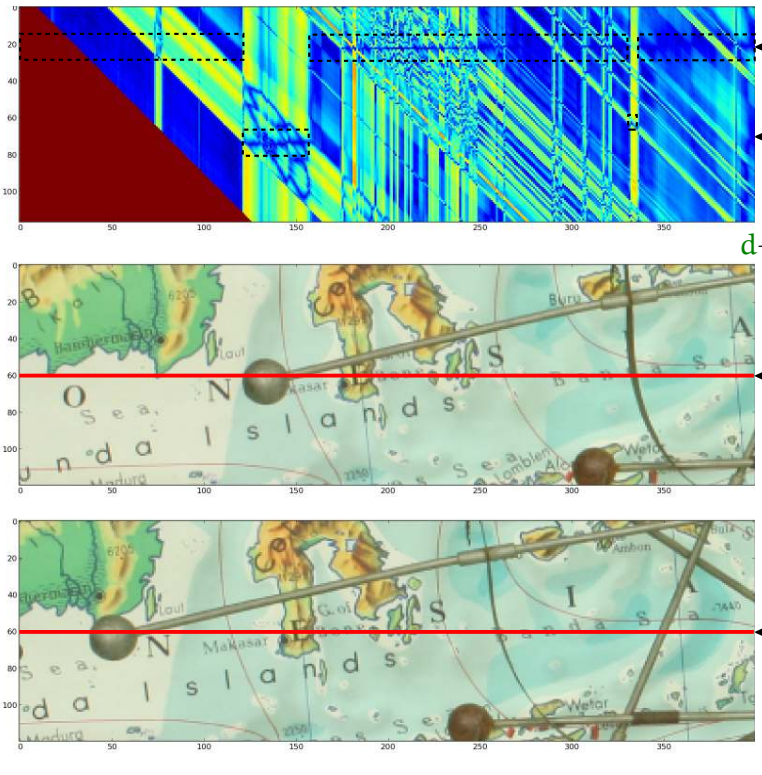
Analysis of Image Superimpositions

Symbol	Description
Stereo images	
$\mathbf{I}_l$	Image associated to the left view
	$\mathbf{I}_l : (x, y) \mapsto \mathbf{I}_l[x, y]$, the intensity of pixel (x, y) in $\mathbf{I}_l$
$\mathbf{I}_r$	Image associated to the right view
	$\mathbf{I}_r : (x, y) \mapsto \mathbf{I}_r[x, y]$, the intensity of pixel (x, y) in $\mathbf{I}_r$

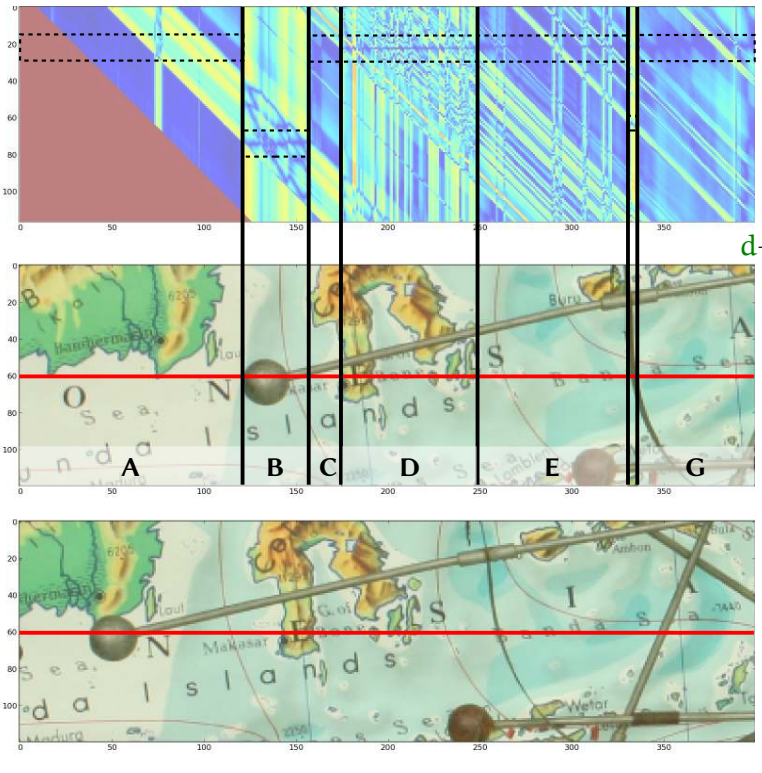
Symbol	Description
Structure encoding the superimpositions	
$\mathbf{D}$	Disparity Space Volume (DSV) based on lightness differences
	$\mathbf{D} : (x, y, d) \mapsto \mathbf{D}[x, y, d] = \lvert\mathbf{I}_l[x, y] - \mathbf{I}_r[x - d, y]\rvert$



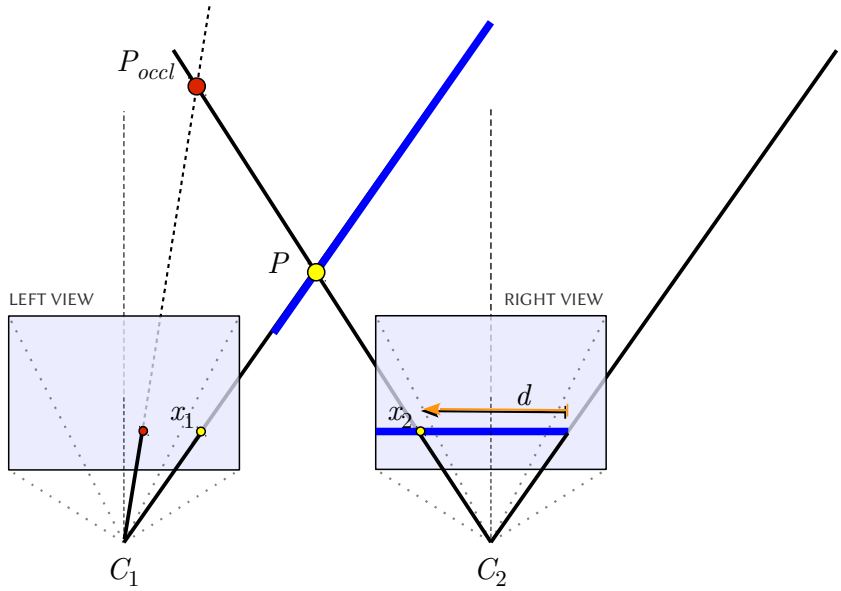
Analysis of Image Superimpositions



Analysis of Image Superimpositions



The Occlusion Phenomenon



The Cross-Checking Criterion

The initial disparity map must take the occlusion phenomenon into account.

Symbol	Description
x_l	Pixel abscissa with respect to the left view
x_r	Pixel abscissa with respect to the right view
$\boldsymbol{W}$	Volume encoding the superimpositions between $\mathbf{I}_l$ and $\mathbf{I}_r$, such that $\boldsymbol{W} : (x_l, x_r, y) \mapsto \boldsymbol{W}[x_l, x_r, y] = \lvert\mathbf{I}_l[x_l, y] - \mathbf{I}_r[x_r, y]\rvert = \mathbf{D}[x_l, y, x_l - x_r]$

Cross-checking ensures that disparity measures are synchronised between left and right views

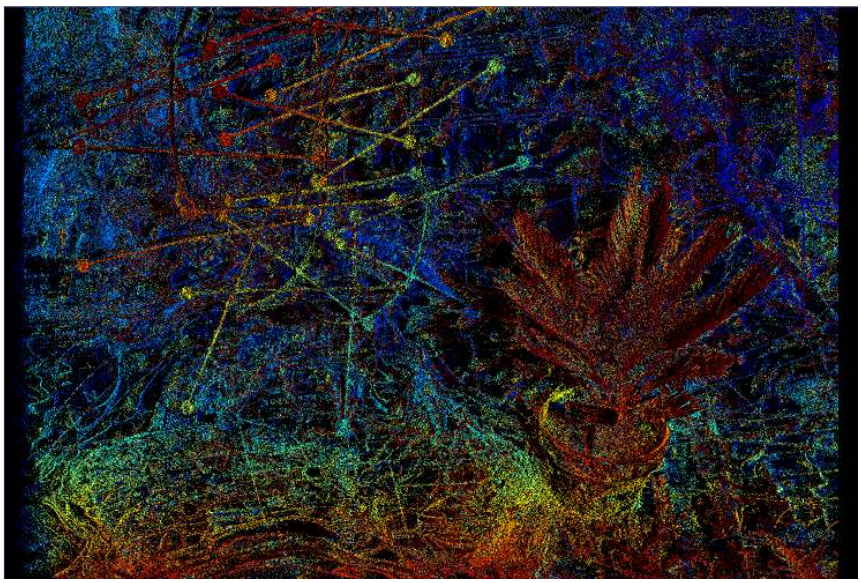
$$\begin{aligned} d &= x_l - x_r \\ \texttt{crossChecking}[x_l, y, d] &= \left(x_r = \arg\min_x \boldsymbol{W}[x_l, x, y]\right) \wedge \left(x_l = \arg\min_x \boldsymbol{W}[x, x_r, y]\right) \end{aligned}$$

➔ Among all the pixels belonging to the right view's scanline of ordinate y, it is (x_r, y) which is best superimposed with (x_l, y) in the left view **and** among all the pixels belonging to the left view's scanline of ordinate y, it is (x_l, y) which is best superimposed with (x_r, y) in the right view.

➔ Points occluded in one of the images of the stereo pair should not satisfy « *cross-checking* »

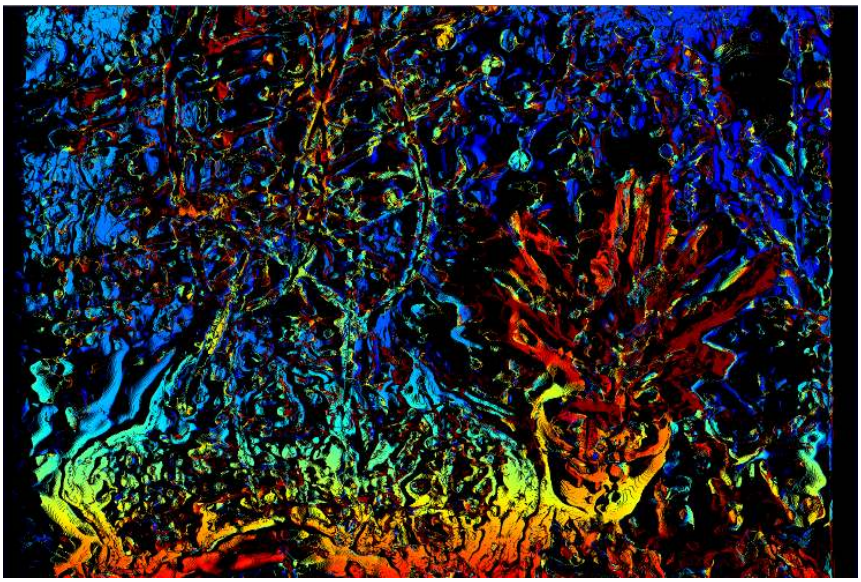
[Fua, 1993]

Initial disparity map deduced from raw DSV



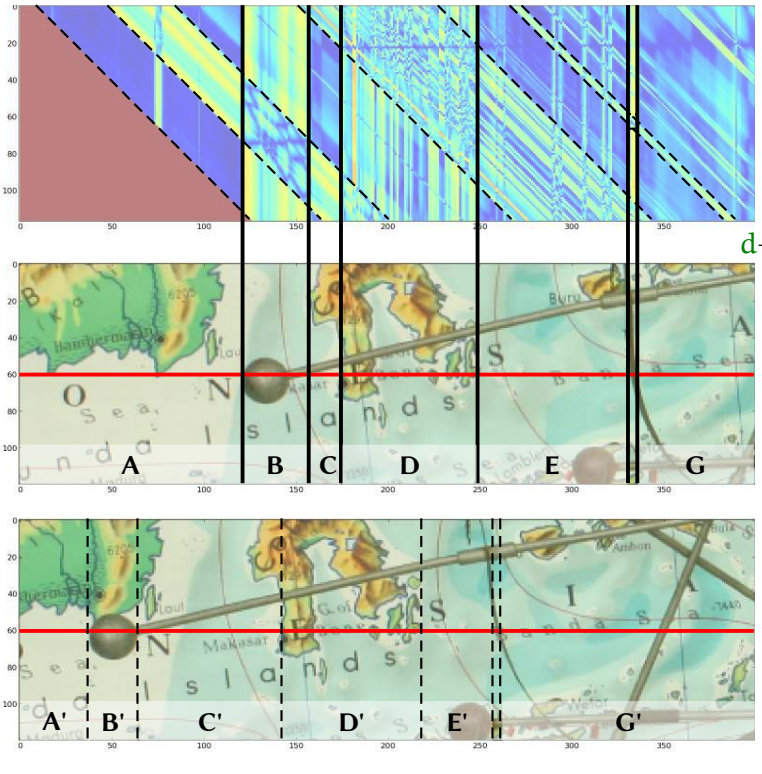
$$\mathfrak{D}^{(\text{init})}[x,y] = \begin{cases} \arg\min_d \mathbf{D}[x,y,d] & \text{if } \texttt{crossChecking}[x,y,d] \text{ is true.} \\ \perp & \text{otherwise} \end{cases}$$

DSV Filtering with a Gaussian Filter



$$\mathbf{D}[x,y,d] \leftarrow \frac{1}{2\pi\sigma^2} \sum_{d_x,d_y \in \{-5,\ldots,+5\}} \mathbf{D}[x+dx, y+dy, d] \cdot e^{-\frac{d_x^2+d_y^2}{2\sigma^2}}$$

DSV Filtering controlled by Segmentation

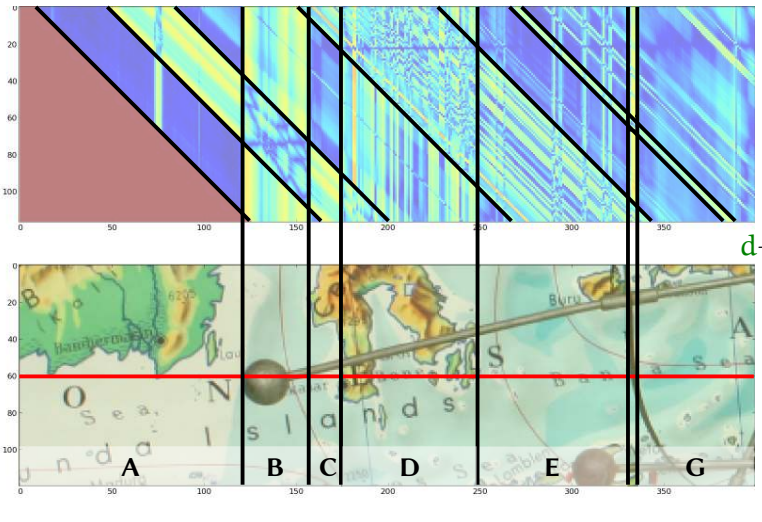


DSV cut along the plane of equation y=60

Left view

Right view

A First Example of Regional Aggregation



DSV cut along the plane of equation y=60

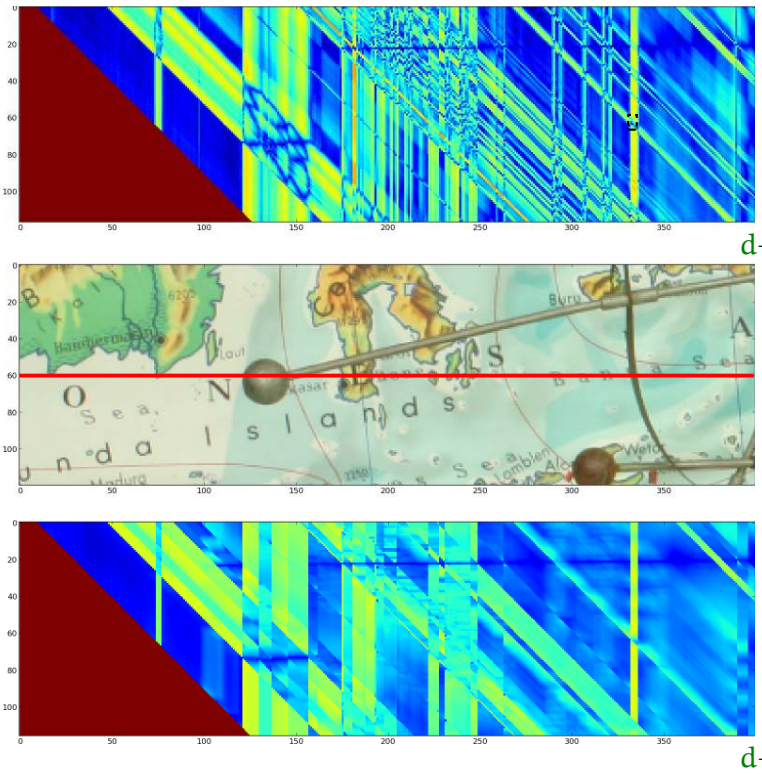
Left view

Suppose we wish to determine which voxels may be used, so that their costs are aggregated to the voxel of coordinates (x,y,d). If (x,y) belongs to region $\mathbf{R}_i$ in the left view, and (x-d,y) belongs to region $\mathbf{R}_j$ in the right view, then :

$$\mathbf{S}_{(x,y,d)} = \{(d_x, d_y) \mid (x + d_x, y + d_y) \in \mathbf{R}_i \wedge (x + d_x - d, y + d_y) \in \mathbf{R}_j\}$$

$$\mathbf{D}[x, y, d] \leftarrow \frac{1}{|\mathbf{S}_{(x,y,d)}|} \sum_{(d_x,d_y)\in\mathbf{S}_{(x,y,d)}} \mathbf{D}[x + d_x, y + d_y, d]$$

A First Example of Regional Aggregation



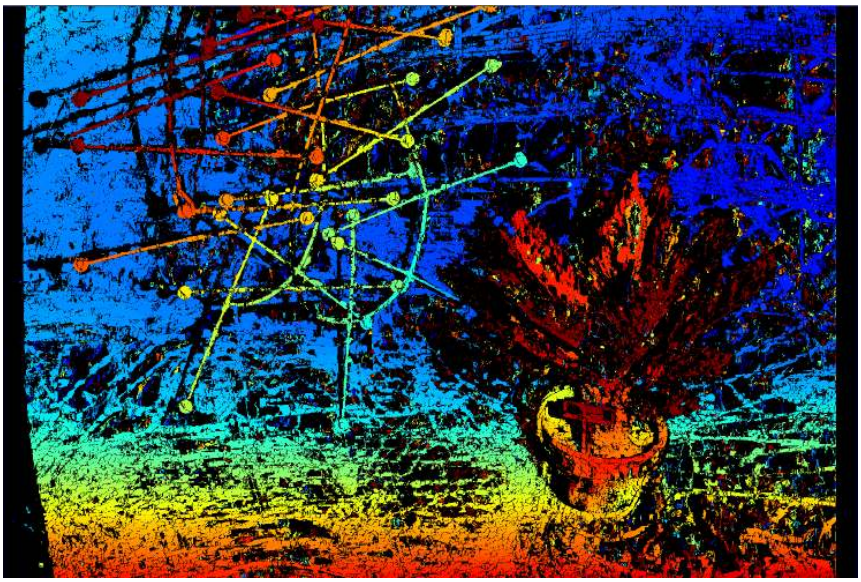
DSV cut along the plane of equation y=60

Left view

FILTERED DSV cut along the plane y=60 with $|d_x|, |d_y| \leq 5$

Dimensions
Patch : 11 x 11 pixels
Image : 1434 x 984 pixels

A First Example of Regional Aggregation



An initial disparity map with
Aggregation constrained by left and right image segmentations

Towards a more elaborate Aggregation System

1. Up to now, the aggregation is controlled by a **patch of limited scope**.
   - ✗ ***Insufficient in texture-less areas***
   - ✗ ***Induces errors and intra-regional inconsistencies***

2. Furthermore, the filtering operates independently on each disparity plane. Therefore, we assumed that the image areas covered by the patches are **fronto-parallel** to the cameras.
   - ✗ ***Little imprecisions as long as patch size is small***
   - ✗ ***Problematic for tilted regions***

Cost Aggregation

Cost Aggregation (Fronto-Parallel)

Consider the horizontal ***moving average***, in the left direction :

$$\tilde{D}_n^{(L)}[x,y,d] = \frac{1}{n+1}\sum_{i=0}^{n} D[x-i,y,d]$$

We split the summation term as follows :

$$\begin{aligned}
\tilde{D}_n^{(L)}[x,y,d] &= \frac{1}{n+1}\left(D[x,y,d] + \sum_{i=1}^{n} D[x-i,y,d]\right) \\
\textit{I set } i = j+1 \quad &= \frac{1}{n+1}\left(D[x,y,d] + \sum_{j=0}^{n-1} D[(x-1)-j,y,d]\right) \\
&= \frac{1}{n+1}\left(D[x,y,d] + n\cdot \tilde{D}_{n-1}^{(L)}[x-1,y,d]\right) \\
\textit{I express the shift as an erosion} \quad &= \frac{1}{n+1}\left(D[x,y,d] + n\cdot \varepsilon_{B_L}\left(\tilde{D}_{n-1}^{(L)}\right)[x,y,d]\right) \\
\tilde{D}_n^{(L)} &= \frac{1}{n+1}\left(D + n\cdot \varepsilon_{B_L}\left(\tilde{D}_{n-1}^{(L)}\right)\right),\ \forall n \geqslant 1
\end{aligned}$$

With structuring element : $B_L = \{(-1,0,0)\}$

Cost Aggregation & Stopping Criterion

Aggregation constrained by segmentation

We must ***interrupt the aggregation*** if we observe a change of region.
How to proceed?

Data :

$\mathcal{L}_l : (x,y) \mapsto \mathcal{L}_l[x,y] = i \quad \text{iff } (x,y) \in R_i$ *Left image partition*

$\mathcal{L}_r : (x,y) \mapsto \mathcal{L}_r[x,y] = j \quad \text{iff } (x,y) \in R'_j$ *Right image partition*

Superimposition volume of partitions : $\mathcal{L}_V$

$$\mathcal{L}_V[x,y,d] = \mathcal{L}_l[x,y] + \mathcal{L}_r[x-d,y]\cdot \max_{x,y} \mathcal{L}_l[x,y]$$

Cost Aggregation & Stopping Criterion

Aggregation constrained by segmentation

We must ***interrupt the aggregation*** if we observe a change of region.
How to proceed?

$\mathcal{L}_V$								
$\varepsilon_B \mathcal{L}_V$								
$\mathbf{D}$	7	7	8	4	3	1	1	0
$\mathcal{D}_B$	18	19	20	**0**	1	2	3	4
$\mathbf{D}_{OUT}$ @ t=0	7	7	8	4	3	1	1	0
$\mathbf{D}_{OUT}$ @ t=1	?	14	15	4	7	4	2	1
$\mathbf{D}_{OUT}$ @ t=2	?	?	22	4	7	8	5	2
$\mathbf{D}_{OUT}$ @ t=3	?	?	?	4	7	8	9	5
$\mathbf{D}_{OUT}$ @ t=4	?	?	?	4	7	8	9	9
$\mathbf{D}_{OUT} \div (\min\{4, \mathcal{D}_B\} + 1)$	?	?	?	**4**	**3.5**	**2.7**	**2.3**	**1.8**

Update law :

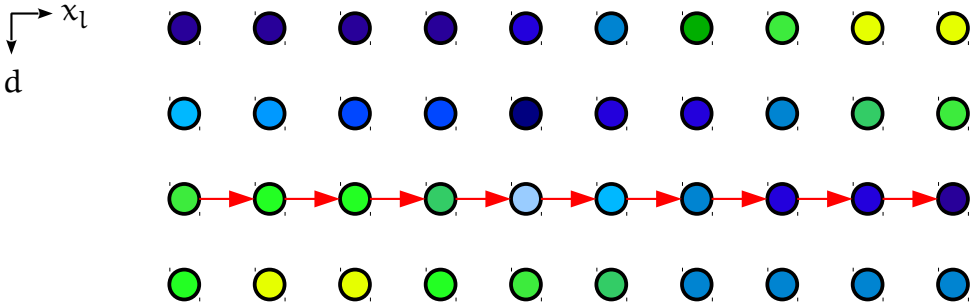
$$\mathbf{D}_{UPD} \leftarrow \mathbf{D} + \varepsilon_B(\mathbf{D}_{OUT})$$

$$\mathbf{D}_{SEL} \leftarrow \text{Binary volume indicating the voxels where } \mathcal{D}_B \geqslant t$$

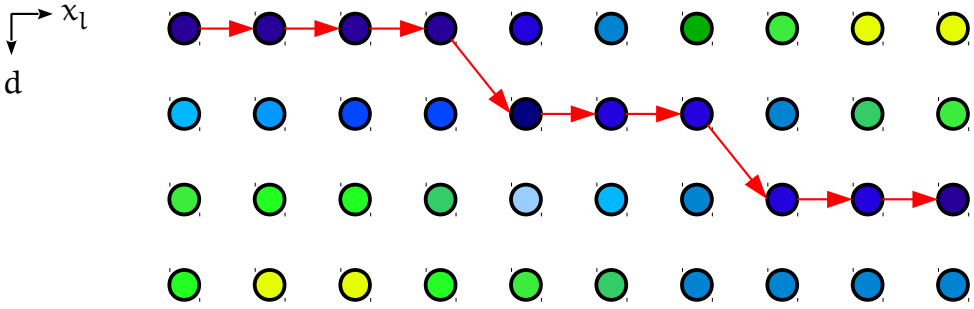
$$\mathbf{D}_{OUT} \leftarrow \mathbf{D}_{UPD} \cdot \mathbf{D}_{SEL} + \mathbf{D}_{OUT} \cdot (1 - \mathbf{D}_{SEL})$$

Cost Aggregation with Disparity Variation

Directional aggregation (fronto-parallel)



Directional aggregation (shortest path)



Cost Aggregation with Disparity Variation

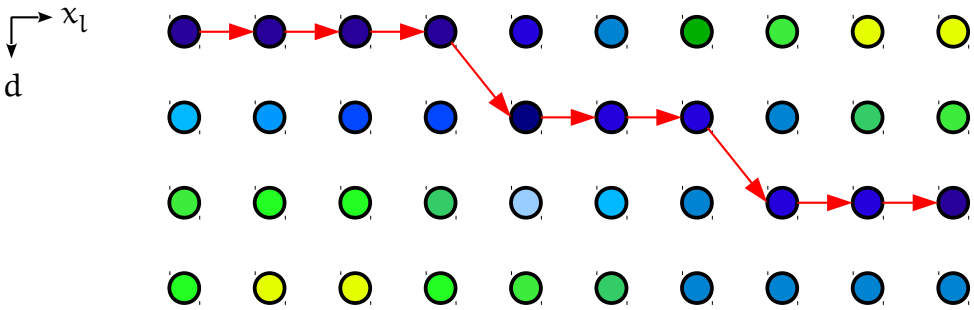
Directional aggregation (shortest path)

regularisation term

$$\tilde{D}_n^{(d)} = \frac{1}{n+1}\left(D + n \cdot \min\left\{\varepsilon_{B_0}\left(\tilde{D}_{n-1}^{(d)}\right), \varepsilon_{B_1}\left(\tilde{D}_{n-1}^{(d)}\right) + \xi\right\}\right), \forall n \geqslant 1$$

Notation :

$d = (d_x, d_y)$		*Image plane direction of diffusion*
$\tilde{D}_0^{(d)} = D$		*Initialisation*
$B_0 = \{(d_x, d_y, 0)\}$		*Fronto-parallel structuring element*
$B_1 = \{(d_x, d_y, -1), (d_x, d_y, +1)\}$		*Structuring element handling tilted regions*

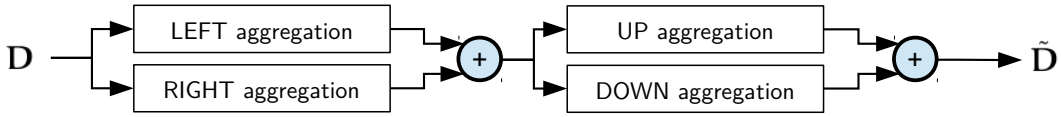


Combining Directional Aggregations

Aggregation directions are expressed with respect to the image plane.
The following directions are considered :

- Horizontal (left and right)
- Vertical (up and down)

The algorithm combines the directional aggregations as follows:

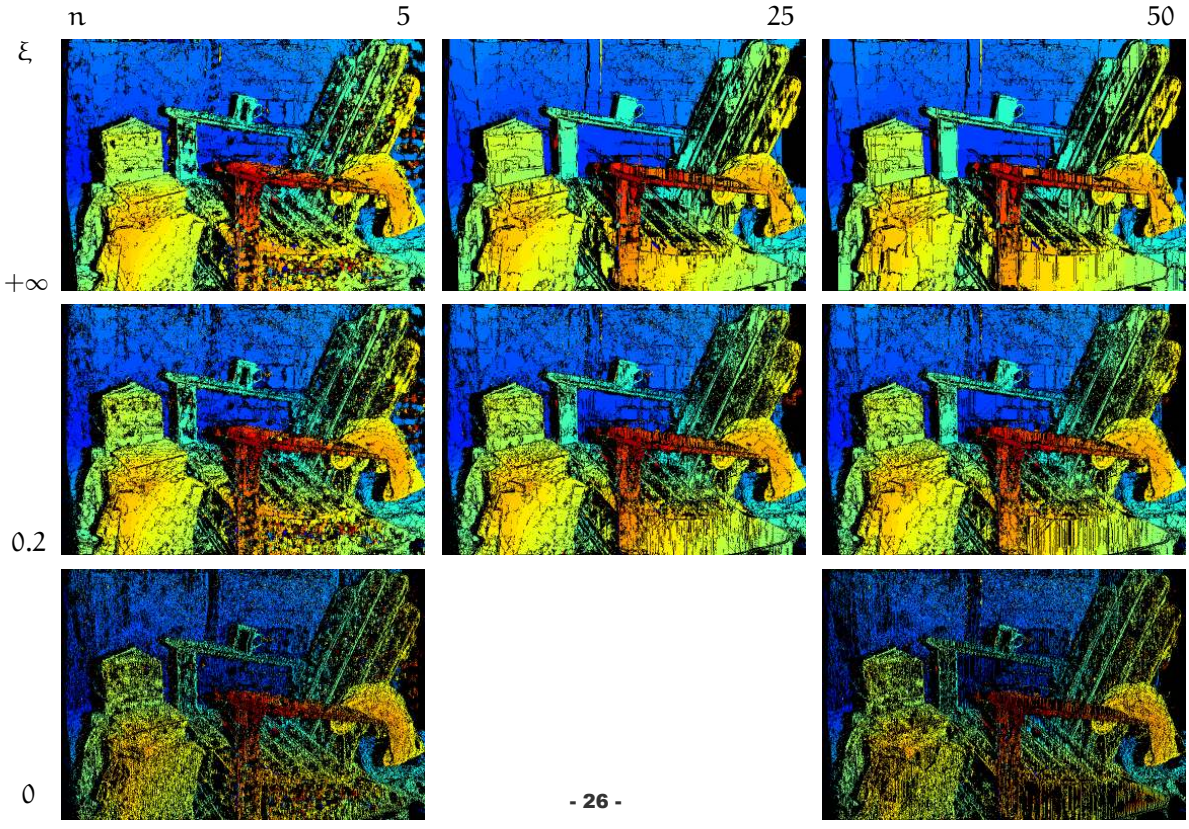


Note :
If there is no disparity variation along the aggregation paths,
then this algorithm implements a 2D moving average (constrained by segmentations),
by exploiting the linear separability of the filter.

Behaviour of the Diffusion Block for different Parameters

Adirondack (Middlebury 2014 database)

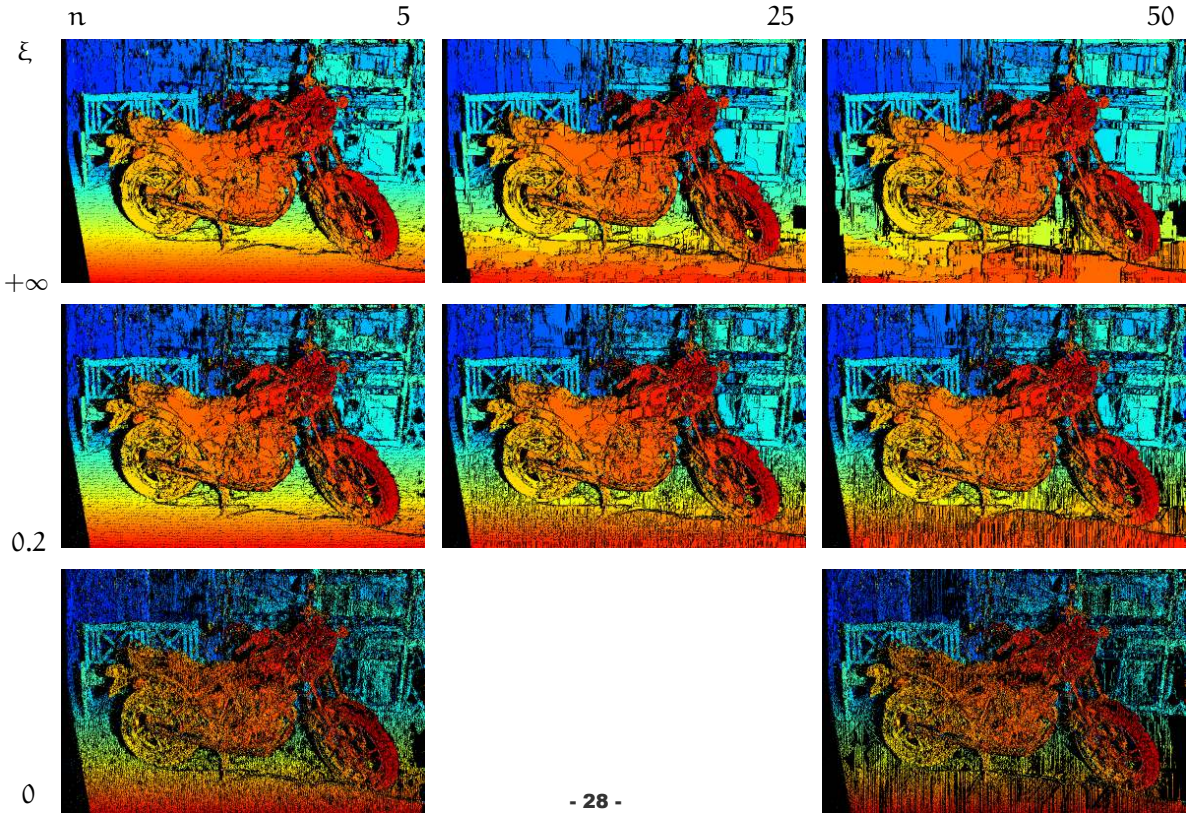
Behaviour of the Diffusion Block for different Parameters



Behaviour of the Diffusion Block for different Parameters

`Motorcycle` (Middlebury 2014 database)

Behaviour of the Diffusion Block for different Parameters

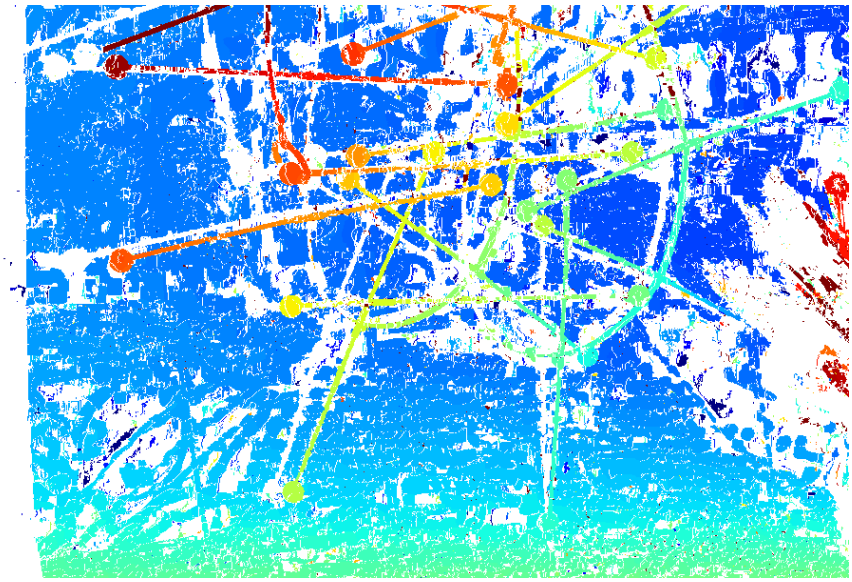


Multi-scale extension to the aggregation system & Filtering of sparse disparity maps

Disparity Map Clustering

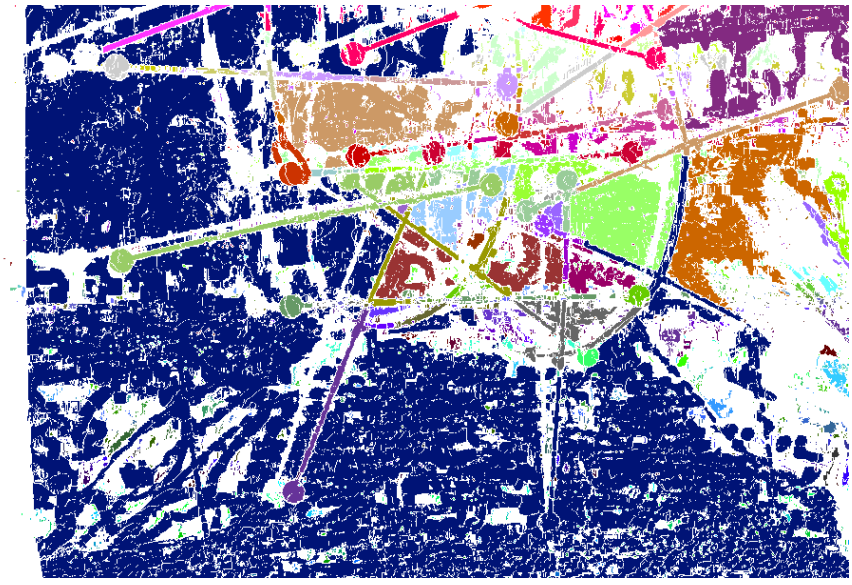
AustraliaP – Initial image

Disparity Map Clustering



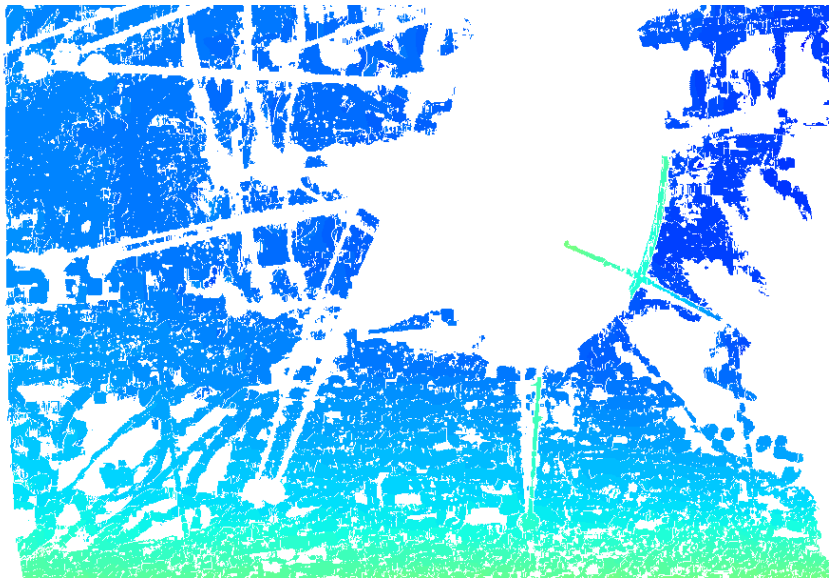
AustraliaP – Initial disparity map (half resolution)

Disparity Map Clustering



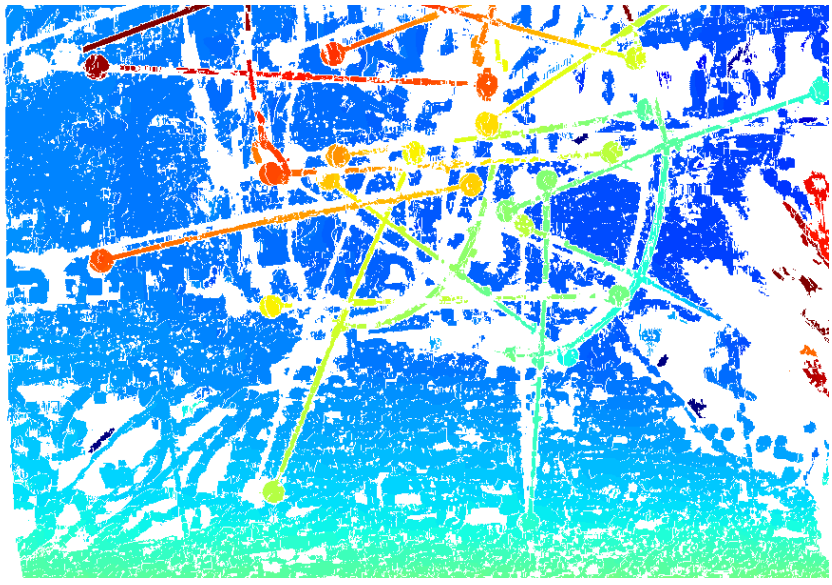
AustraliaP – Disparity clusters

Disparity Map Clustering



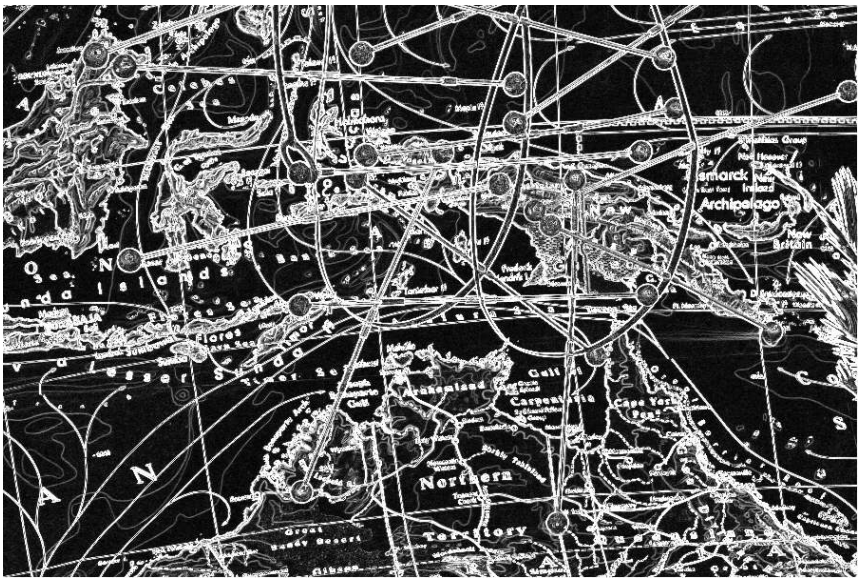
AustraliaP – Large area opening on disparity clusters

Disparity Map Clustering



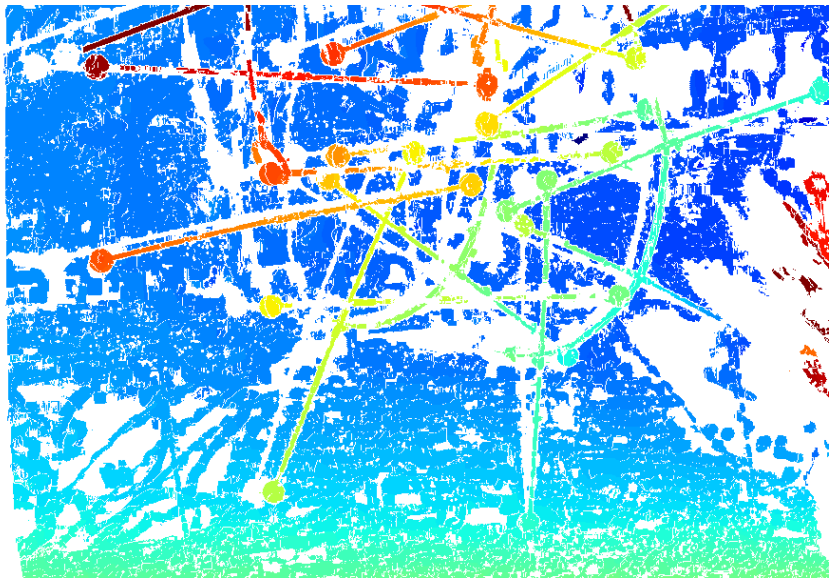
AustraliaP – Smaller area opening on clusters

Disparity Map Clustering



`AustraliaP` – Bad clusters cover low gradient areas

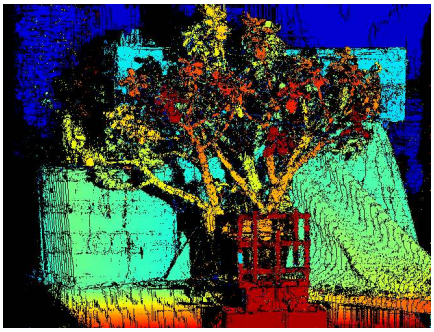
Filtering with Disparity Map Clusters

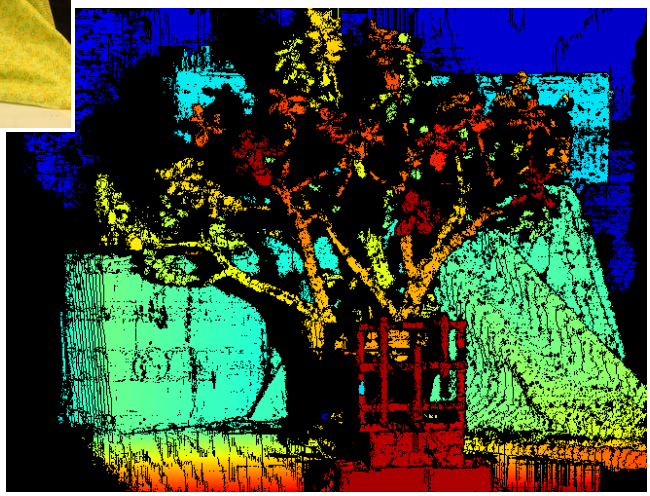


`AustraliaP` – Large clusters, and small clusters holding sufficient gradient information

Filtering with Disparity Map Clusters

RESULT AFTER CLUSTER-BASED FILTERING



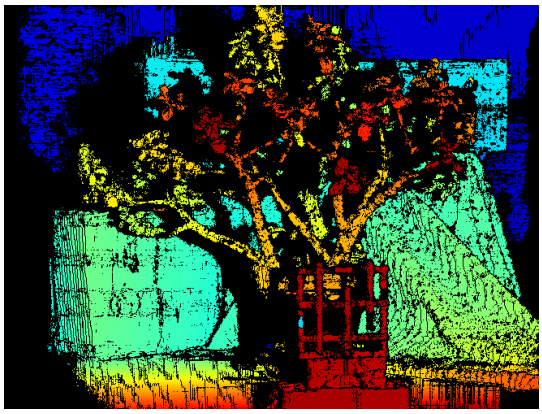


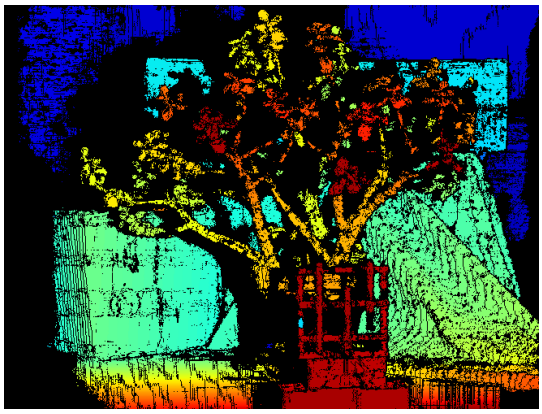
INITIAL RESULT

Fattening Effect Removal

Algorithm's main principles

1. Compute the following elements :
   (a) Disparity clusters on a regional basis
   (b) The first percentiles of the disparity measures available in each region.
   In our experiments, we go up to the 30^{th} percentile.
2. In each region, mark the pixels having a disparity measure equal to one of the percentiles computed for the considered region.
3. Reconstruct clusters holding marked pixels, with their disparities.





BEFORE

AFTER

Multi-Scale Aggregation

Some observations :

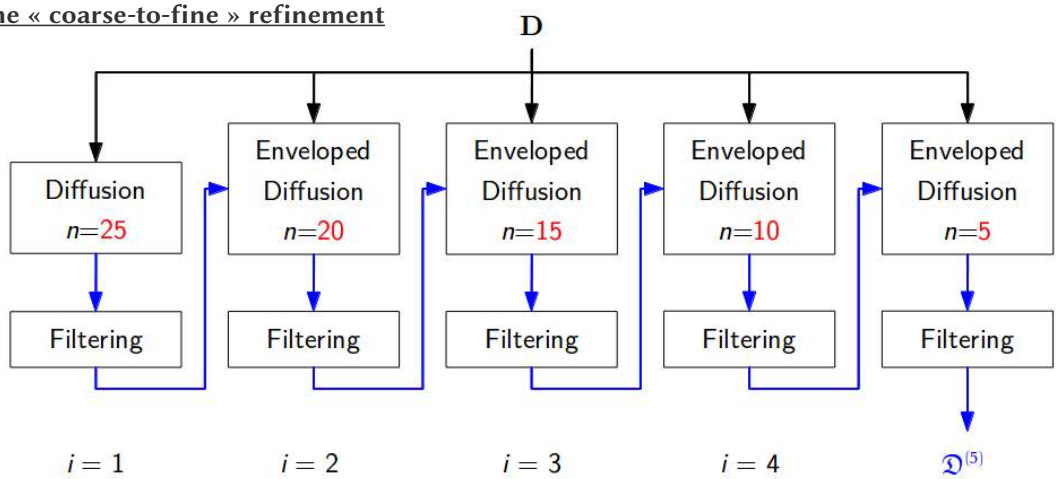
- We now have filtered disparity maps, but relatively sparse.
- The density of measures found in these disparity maps is about 45 % on average.
- Above n=25 pixels, augmenting the aggregation scope neither increases the density of the measures, nor improves the results.
- A small aggregation scope leads to precise measures, but also very gross errors.
- A large aggregation scope reduces the number of gross errors, but yields more approximate disparity measures.

We propose a multi-scale extension in two steps :

- A « ***coarse-to-fine*** » refinement, which yields precise disparity measures, while preserving the regional consistency observed at the coarse level of aggregation.
- A « ***fine-to-coarse*** » densification, which leaves only the pixels found in occluded areas, without a disparity measure.

Multi-Scale Aggregation

The « coarse-to-fine » refinement

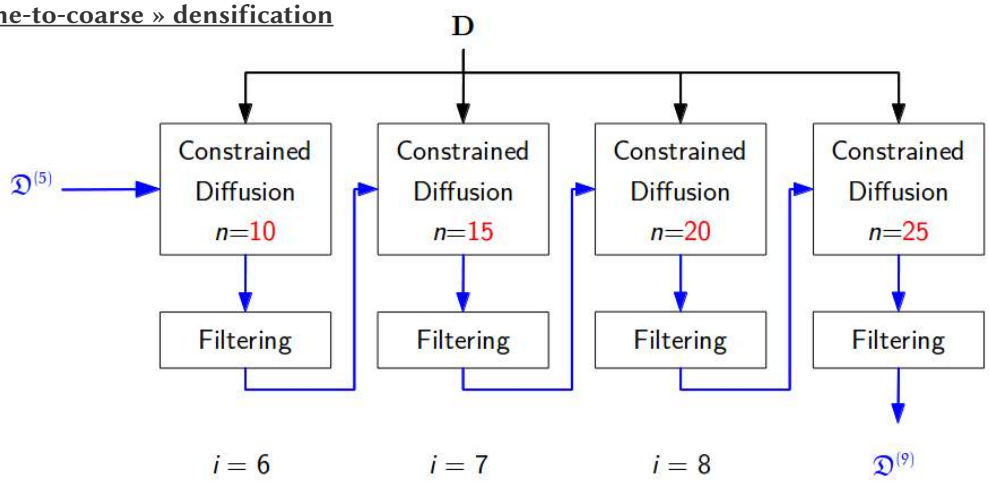


We impose a constraint on the filtered disparity space volume which, at iteration i+1, is:

$$\mathbf{D}_{\mathrm{OUT}}[x,y,d] \leftarrow \begin{cases} \mathbf{D}_{\mathrm{OUT}}[x,y,d] & \text{if } |\mathfrak{D}^{(i)}[x,y] - d| < \Delta^{(i)} \\ +\infty & \text{otherwise} \end{cases}$$

Multi-Scale Aggregation

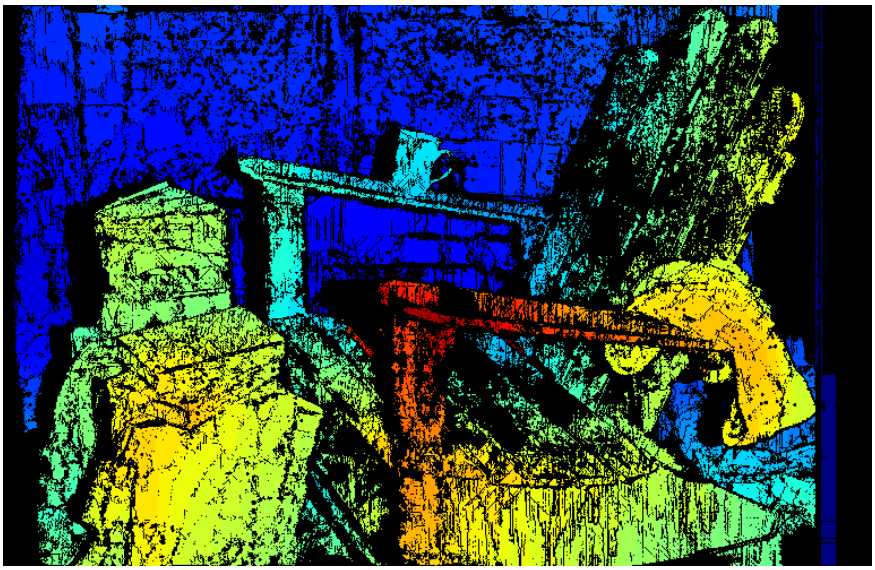
The« fine-to-coarse » densification



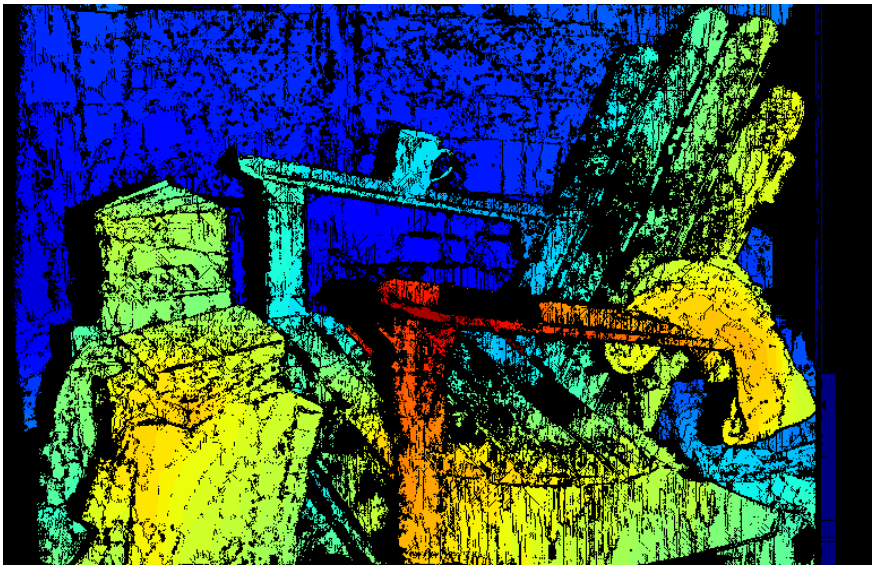
We impose a constraint on the filtered disparity space volume which, at iteration i+1, is:

$$\mathbf{D}_{OUT}[x, y, d] \leftarrow \begin{cases} \mathbf{D}_{OUT}[x, y, d] & \text{if } \mathfrak{D}^{(i)}[x, y] \neq d \\ 0 & \text{otherwise} \end{cases}$$

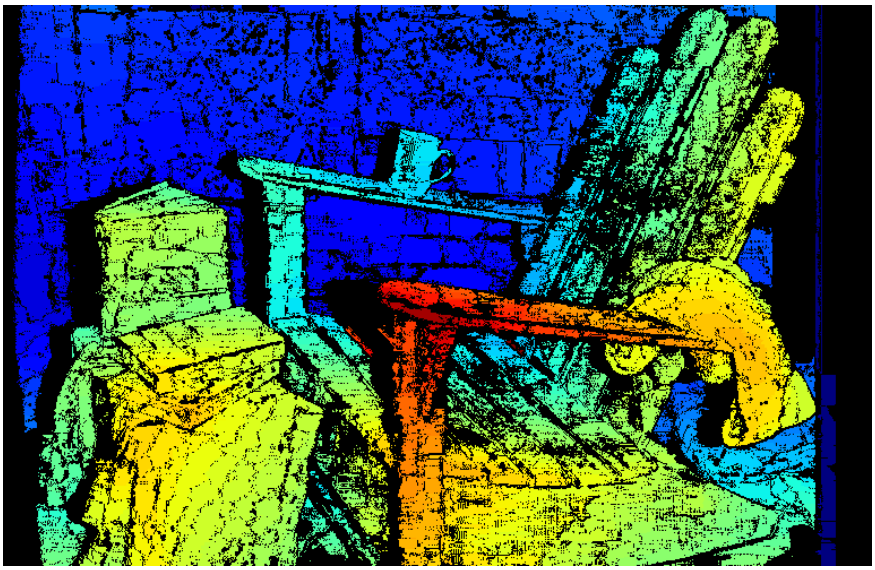
Multi-Scale Aggregation



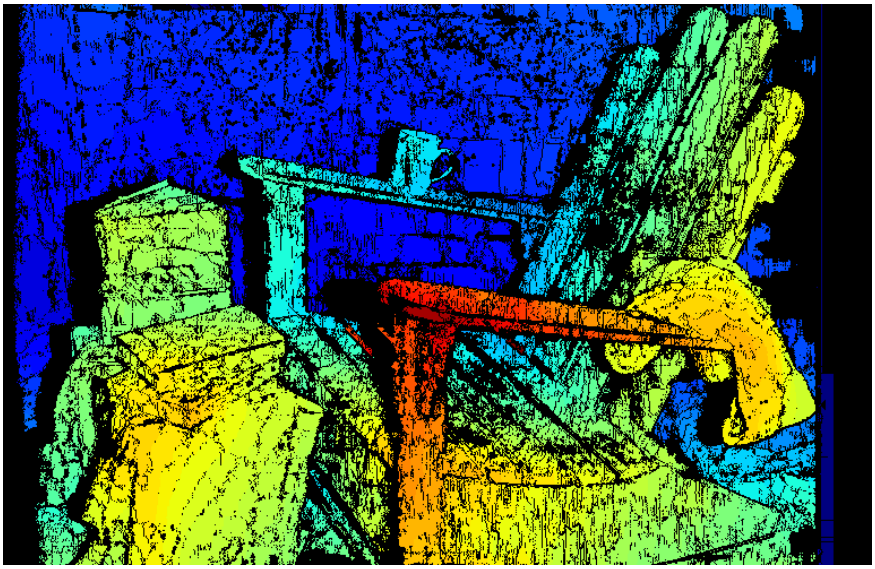
`Adirondack` (Standard diffusion, n=25)



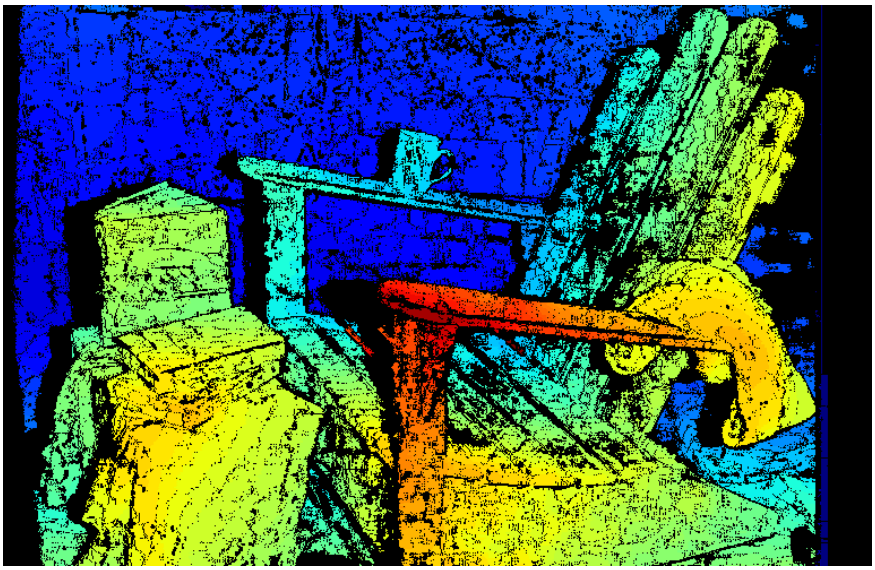
`Adirondack` (Refinement, n=20)



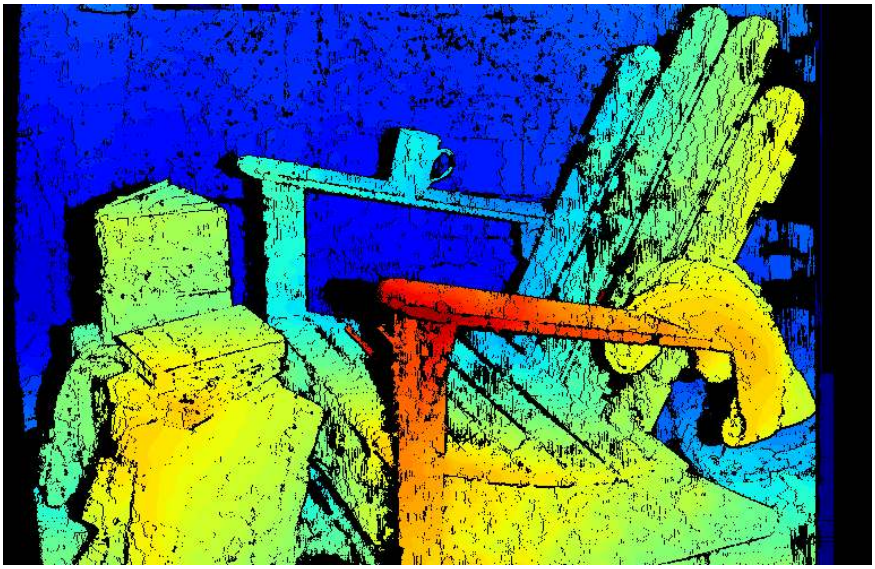
`Adirondack` (Refinement, n=15)



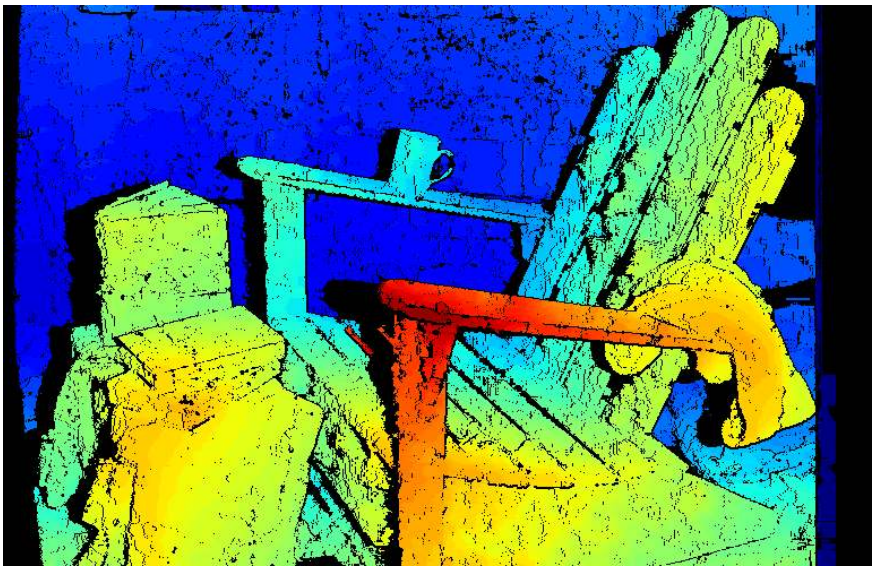
`Adirondack` (Refinement, n=10)



`Adirondack` (Refinement, n=5)

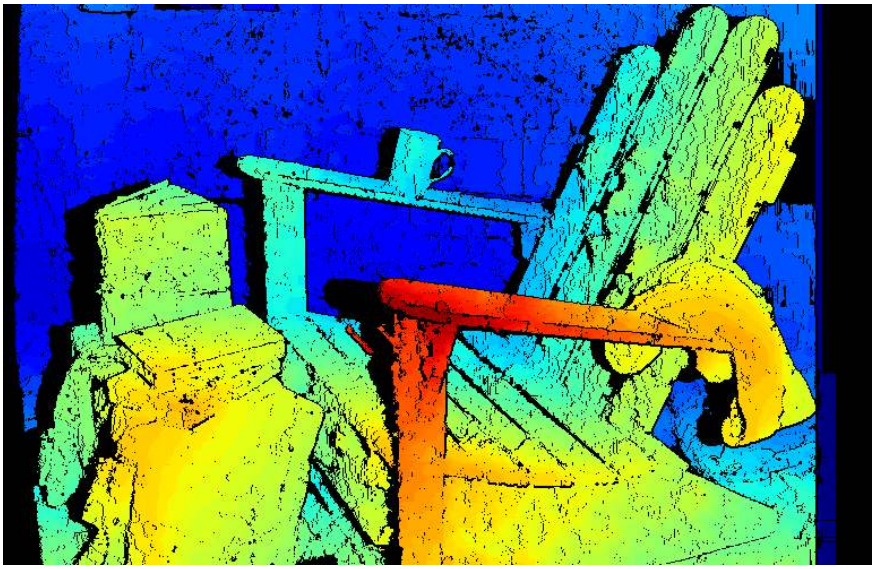


`Adirondack` (Densification, n=10)



`Adirondack` (Densification, n=15)

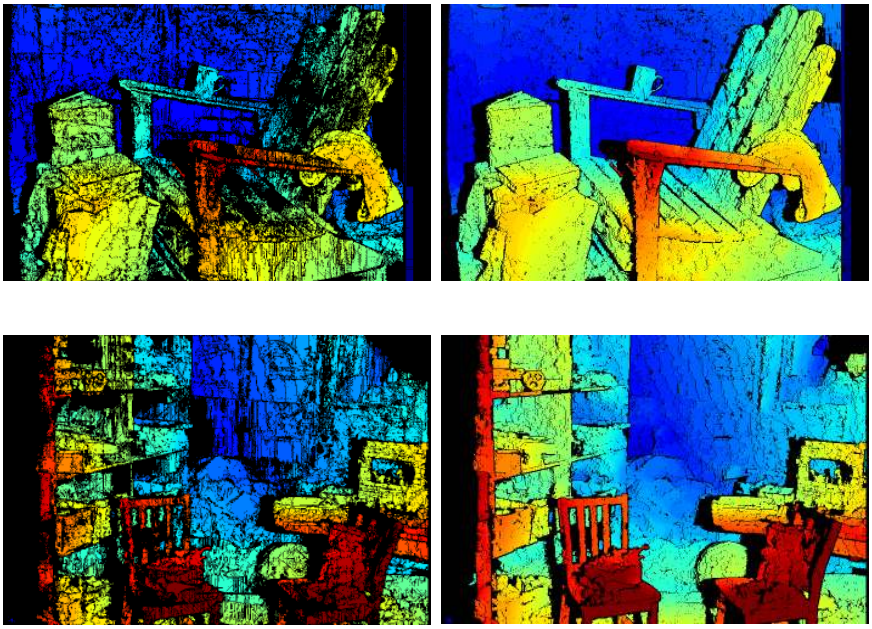
Multi-Scale Aggregation



`Adirondack` (Densification, n=20)

Multi-Scale Aggregation

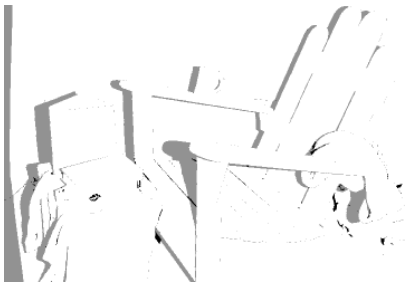
INITIAL DIFFUSION MULTI-SCALE DIFFUSION

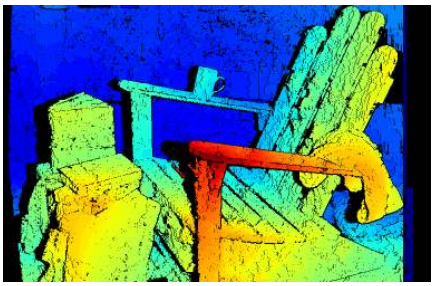


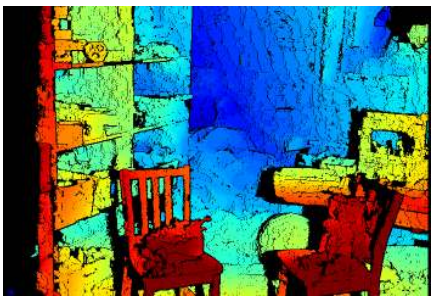
Multi-Scale Aggregation

OCCLUSION MASKS

MULTI-SCALE DIFFUSION



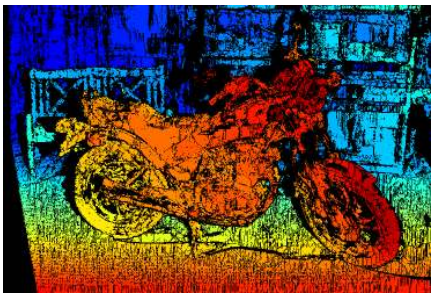


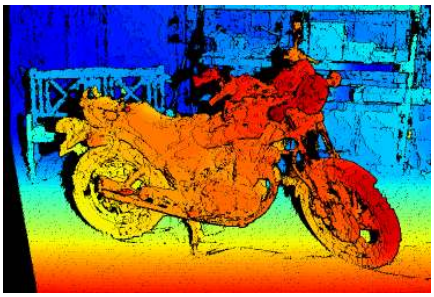


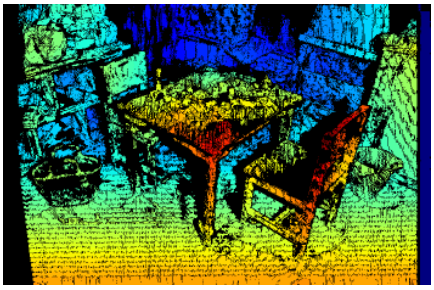
Multi-Scale Aggregation

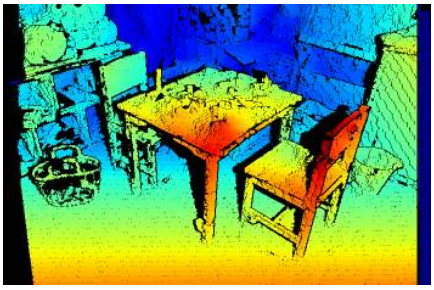
INITIAL DIFFUSION

MULTI-SCALE DIFFUSION



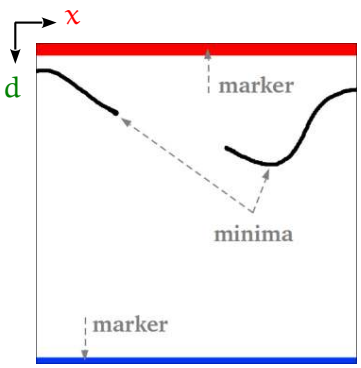




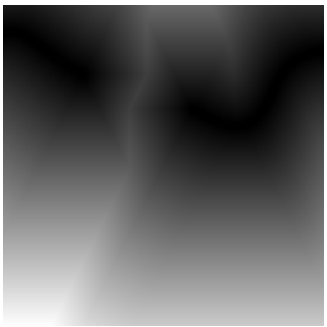


Hole filling and Finishing

Interpolation, Distance Functions, and Watershed Transformation

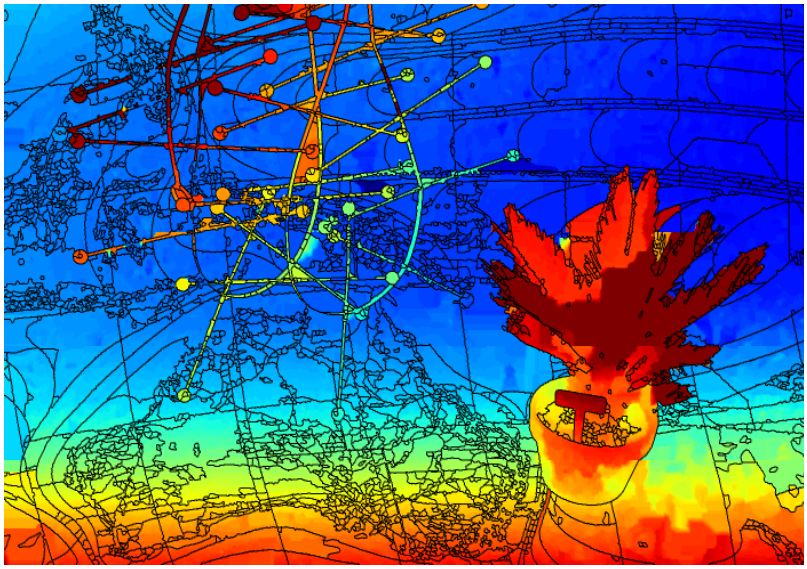


Configuration

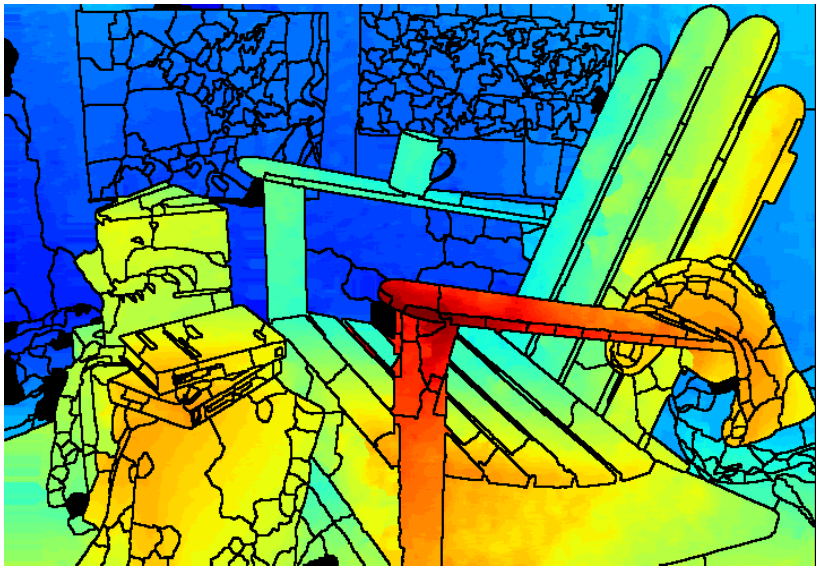


Binary distance function obtained by successive erosions on minima

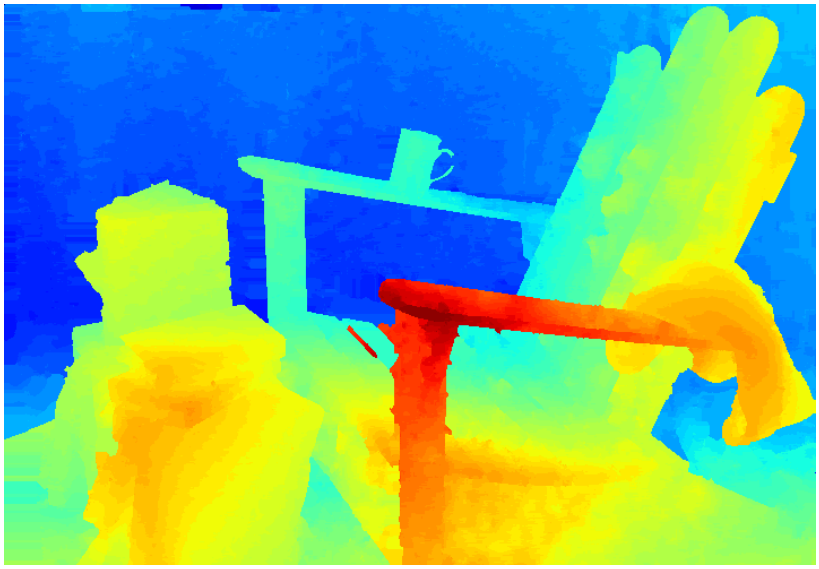
Watershed of the inverted distance function, controlled by the red and blue markers



`AustraliaP` (Interpolation based on distance functions, per cell)

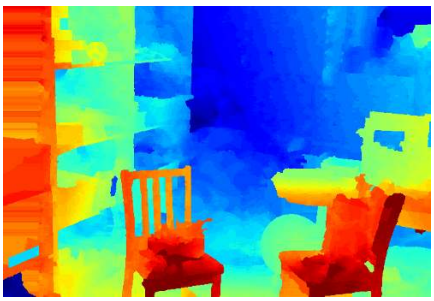


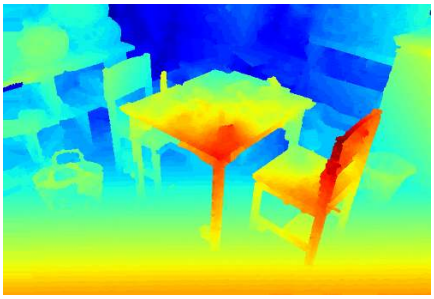
`Adirondack` (Interpolation based on distance functions, per cell)



Adirondack (Finishing)

REFERENCE VIEW FULL DISPARITY MAP





Evaluation

Middlebury 2014 Benchmark

Online evaluation system : ***http://vision.middlebury.edu/stereo***

- 30 evaluated stereo pairs, with **resolution choice** :
  - ➔ Full [F] – 2872 x 1984 pixels
  - ➔ Half [H] – 1436 x 992 pixels
  - ➔ Quarter [Q] – 718 x 496 pixels (chosen)

- **Wide-baseline** (200 levels of disparity for resolution [Q], 800 levels for resolution [F])

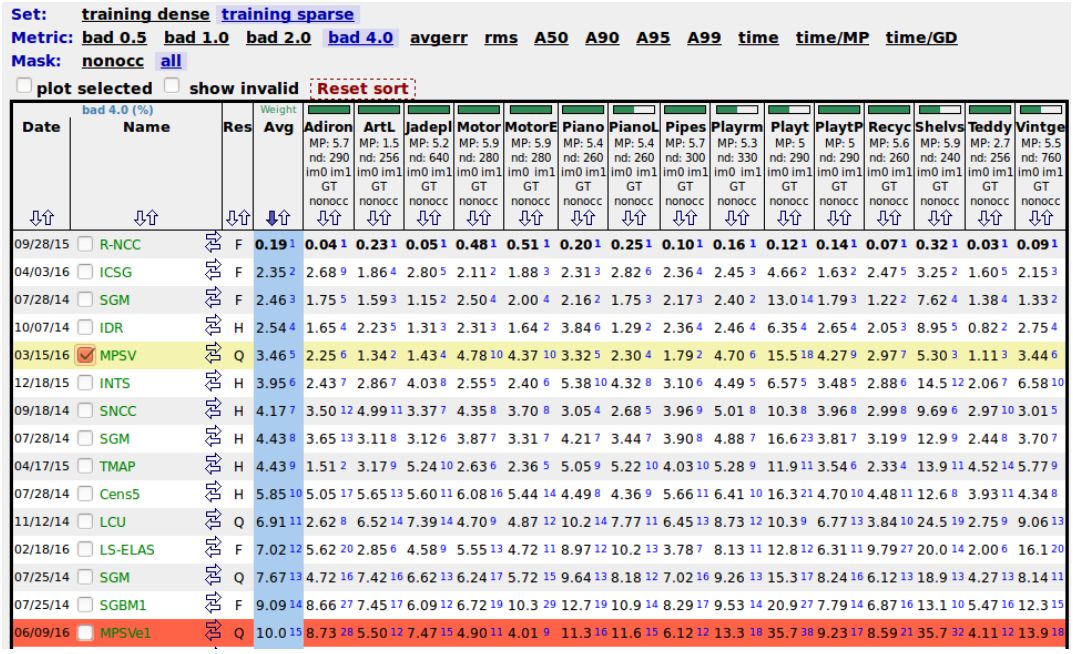
- Accuracy measures (examples) :
  - ➔ « **Bad pixels 4** » : percentage of pixels with a disparity measure differing from 4 pixels or more, with respect to the ground truth at full resolution. At quarter resolution, the maximum disparity error threshold equals 1 pixel.
  - ➔ « **Average error** » : the mean average error compared to the ground truth. (expressed according to resolution [F])
  - ➔ « **RMS error** » : the root-mean-square error compared to the ground truth. (expressed according to resolution [F])

- **Evaluation performed for our sparse and full disparity maps.**
- **Important: the level of densification must be taken into account for the sparse evaluation.**

MIDDLEBURY 2014 BENCHMARK

Sparse results (compared to 44 methods)

MPSV = Diffusion + Filtering, **MPSVe1** = Multi-scale diffusion + Filtering



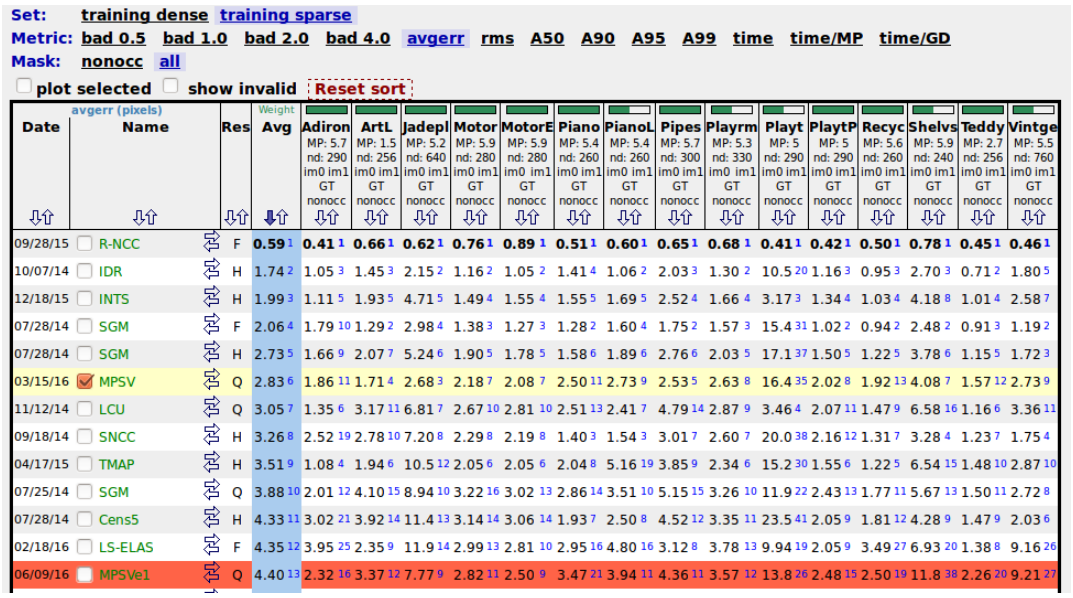
Set: training dense training sparse
Metric: bad 0.5 bad 1.0 bad 2.0 bad 4.0 avgerr rms A50 A90 A95 A99 time time/MP time/GD
Mask: nonocc all
plot selected show invalid Reset sort

Date	Name	Res	Avg (bad 4.0 (%))	Adiron MP: 5.7 nd: 290 im0 im1 GT nonocc	ArtL MP: 1.5 nd: 256 im0 im1 GT nonocc	Jadepl MP: 5.2 nd: 640 im0 im1 GT nonocc	Motor MP: 5.9 nd: 280 im0 im1 GT nonocc	MotorE MP: 5.9 nd: 280 im0 im1 GT nonocc	Piano MP: 5.4 nd: 260 im0 im1 GT nonocc	PianoL MP: 5.4 nd: 260 im0 im1 GT nonocc	Pipes MP: 5.7 nd: 300 im0 im1 GT nonocc	Playrm MP: 5.3 nd: 330 im0 im1 GT nonocc	Playt MP: 5 nd: 290 im0 im1 GT nonocc	PlaytP MP: 5 nd: 290 im0 im1 GT nonocc	Recyc MP: 5.6 nd: 260 im0 im1 GT nonocc	Shelvs MP: 5.9 nd: 240 im0 im1 GT nonocc	Teddy MP: 2.7 nd: 256 im0 im1 GT nonocc	Vintge MP: 5.5 nd: 760 im0 im1 GT nonocc
09/28/15	R-NCC	F	**0.19** 1	**0.04** 1	**0.23** 1	**0.05** 1	**0.48** 1	**0.51** 1	**0.20** 1	**0.25** 1	**0.10** 1	**0.16** 1	**0.12** 1	**0.14** 1	**0.07** 1	**0.32** 1	**0.03** 1	**0.09** 1
04/03/16	ICSG	F	2.35 2	2.68 9	1.86 4	2.80 5	2.11 2	1.88 3	2.31 3	2.82 6	2.36 4	2.45 3	4.66 2	1.63 2	2.47 5	3.25 2	1.60 5	2.15 3
07/28/14	SGM	F	2.46 3	1.75 5	1.59 3	1.15 2	2.50 4	2.00 4	2.16 2	1.75 3	2.17 3	2.40 2	13.0 14	1.79 3	1.22 2	7.62 4	1.38 4	1.33 2
10/07/14	IDR	H	2.54 4	1.65 4	2.23 5	1.31 3	2.31 3	1.64 2	3.84 6	1.29 2	2.36 4	2.46 4	6.35 4	2.65 4	2.05 3	8.95 5	0.82 2	2.75 4
03/15/16	MPSV	Q	3.46 5	2.25 6	1.34 2	1.43 4	4.78 10	4.37 10	3.32 5	2.30 4	1.79 2	4.70 6	15.5 18	4.27 9	2.97 7	5.30 3	1.11 3	3.44 6
12/18/15	INTS	H	3.95 6	2.43 7	2.86 7	4.03 8	2.55 5	2.40 6	5.38 10	4.32 8	3.10 6	4.49 5	6.57 5	3.48 5	2.88 6	14.5 12	2.06 7	6.58 10
09/18/14	SNCC	H	4.17 7	3.50 12	4.99 11	3.37 7	4.35 8	3.70 8	3.05 4	2.68 5	3.96 9	5.01 8	10.3 8	3.96 8	2.99 8	9.69 6	2.97 10	3.01 5
07/28/14	SGM	H	4.43 8	3.65 13	3.11 8	3.12 6	3.87 7	3.31 7	4.21 7	3.44 7	3.90 8	4.88 7	16.6 23	3.81 7	3.19 9	12.9 9	2.44 8	3.70 7
04/17/15	TMAP	H	4.43 9	1.51 2	3.17 9	5.24 10	2.63 6	2.36 5	5.05 9	5.22 10	4.03 10	5.28 9	11.9 11	3.54 6	2.33 4	13.9 11	4.52 14	5.77 9
07/28/14	Cens5	H	5.85 10	5.05 17	5.65 13	5.60 11	6.08 16	5.44 14	4.49 8	4.36 9	5.66 11	6.41 10	16.3 21	4.70 10	4.48 11	12.6 8	3.93 11	4.34 8
11/12/14	LCU	Q	6.91 11	2.62 8	6.52 14	7.39 14	4.70 9	4.87 12	10.2 14	7.77 11	6.45 13	8.73 12	10.3 9	6.77 13	3.84 10	24.5 19	2.75 9	9.06 13
02/18/16	LS-ELAS	F	7.02 12	5.62 20	2.85 6	4.58 9	5.55 13	4.72 11	8.97 12	10.2 13	3.78 7	8.13 11	12.8 12	6.31 11	9.79 27	20.0 14	2.00 6	16.1 20
07/25/14	SGM	Q	7.67 13	4.72 16	7.42 16	6.62 13	6.24 17	5.72 15	9.64 13	8.18 12	7.02 16	9.26 13	15.3 17	8.24 16	6.12 13	18.9 13	4.27 13	8.14 11
07/25/14	SGBM1	F	9.09 14	8.66 27	7.45 17	6.09 12	6.72 19	10.3 29	12.7 19	10.9 14	8.29 17	9.53 14	20.9 27	7.79 14	6.87 16	13.1 10	5.47 16	12.3 15
06/09/16	MPSVe1	Q	10.0 15	8.73 28	5.50 12	7.47 15	4.90 11	4.01 9	11.3 16	11.6 15	6.12 12	13.3 18	35.7 38	9.23 17	8.59 21	35.7 32	4.11 12	13.9 18

MIDDLEBURY 2014 BENCHMARK

Sparse results (compared to 44 methods)

MPSV = Diffusion + Filtering, **MPSVe1** = Multi-scale diffusion + Filtering



Set: training dense training sparse
Metric: bad 0.5 bad 1.0 bad 2.0 bad 4.0 avgerr rms A50 A90 A95 A99 time time/MP time/GD
Mask: nonocc all
plot selected show invalid Reset sort

Date	Name	Res	Avg (avgerr (pixels))	Adiron MP: 5.7 nd: 290 im0 im1 GT nonocc	ArtL MP: 1.5 nd: 256 im0 im1 GT nonocc	Jadepl MP: 5.2 nd: 640 im0 im1 GT nonocc	Motor MP: 5.9 nd: 280 im0 im1 GT nonocc	MotorE MP: 5.9 nd: 280 im0 im1 GT nonocc	Piano MP: 5.4 nd: 260 im0 im1 GT nonocc	PianoL MP: 5.4 nd: 260 im0 im1 GT nonocc	Pipes MP: 5.7 nd: 300 im0 im1 GT nonocc	Playrm MP: 5.3 nd: 330 im0 im1 GT nonocc	Playt MP: 5 nd: 290 im0 im1 GT nonocc	PlaytP MP: 5 nd: 290 im0 im1 GT nonocc	Recyc MP: 5.6 nd: 260 im0 im1 GT nonocc	Shelvs MP: 5.9 nd: 240 im0 im1 GT nonocc	Teddy MP: 2.7 nd: 256 im0 im1 GT nonocc	Vintge MP: 5.5 nd: 760 im0 im1 GT nonocc
09/28/15	R-NCC	F	**0.59** 1	**0.41** 1	**0.66** 1	**0.62** 1	**0.76** 1	**0.89** 1	**0.51** 1	**0.60** 1	**0.65** 1	**0.68** 1	**0.41** 1	**0.42** 1	**0.50** 1	**0.78** 1	**0.45** 1	**0.46** 1
10/07/14	IDR	H	1.74 2	1.05 3	1.45 3	2.15 2	1.16 2	1.05 2	1.41 4	1.06 2	2.03 3	1.30 2	10.5 20	1.16 3	0.95 3	2.70 3	0.71 2	1.80 5
12/18/15	INTS	H	1.99 3	1.11 5	1.93 5	4.71 5	1.49 4	1.55 4	1.55 5	1.69 5	2.52 4	1.66 4	3.17 3	1.34 4	1.03 4	4.18 8	1.01 4	2.58 7
07/28/14	SGM	F	2.06 4	1.79 10	1.29 2	2.98 4	1.38 3	1.27 3	1.28 2	1.60 4	1.75 2	1.57 3	15.4 31	1.02 2	0.94 2	2.48 2	0.91 3	1.19 2
07/28/14	SGM	H	2.73 5	1.66 9	2.07 7	5.24 6	1.90 5	1.78 5	1.58 6	1.89 6	2.76 6	2.03 5	17.1 37	1.50 5	1.22 5	3.78 6	1.15 5	1.72 3
03/15/16	MPSV	Q	2.83 6	1.86 11	1.71 4	2.68 3	2.18 7	2.08 7	2.50 11	2.73 9	2.53 5	2.63 8	16.4 35	2.02 8	1.92 13	4.08 7	1.57 12	2.73 9
11/12/14	LCU	Q	3.05 7	1.35 6	3.17 11	6.81 7	2.67 10	2.81 10	2.51 13	2.41 7	4.79 14	2.87 9	3.46 4	2.07 11	1.47 9	6.58 16	1.16 6	3.36 11
09/18/14	SNCC	H	3.26 8	2.52 19	2.78 10	7.20 8	2.29 8	2.19 8	1.40 3	1.54 3	3.01 7	2.60 7	20.0 38	2.16 12	1.31 7	3.28 4	1.23 7	1.75 4
04/17/15	TMAP	H	3.51 9	1.08 4	1.94 6	10.5 12	2.05 6	2.05 6	2.04 8	5.16 19	3.85 9	2.34 6	15.2 30	1.55 6	1.22 5	6.54 15	1.48 10	2.87 10
07/25/14	SGM	Q	3.88 10	2.01 12	4.10 15	8.94 10	3.22 16	3.02 13	2.86 14	3.51 10	5.15 15	3.26 10	11.9 22	2.43 13	1.77 11	5.67 13	1.50 11	2.72 8
07/28/14	Cens5	H	4.33 11	3.02 21	3.92 14	11.4 13	3.14 14	3.06 14	1.93 7	2.50 8	4.52 12	3.35 11	23.5 41	2.05 9	1.81 12	4.28 9	1.47 9	2.03 6
02/18/16	LS-ELAS	F	4.35 12	3.95 25	2.35 9	11.9 14	2.99 13	2.81 10	2.95 16	4.80 16	3.12 8	3.78 13	9.94 19	2.05 9	3.49 27	6.93 20	1.38 8	9.16 26
06/09/16	MPSVe1	Q	4.40 13	2.32 16	3.37 12	7.77 9	2.82 11	2.50 9	3.47 21	3.94 11	4.36 11	3.57 12	13.8 26	2.48 15	2.50 19	11.8 38	2.26 20	9.21 27

Sparse results (compared to 44 methods)

MPSV = Diffusion + Filtering, **MPSVe1** = Multi-scale diffusion + Filtering

Set: training dense training sparse
Metric: bad 0.5 bad 1.0 bad 2.0 bad 4.0 avgerr rms A50 A90 A95 A99 time time/MP time/GD
Mask: nonocc all
plot selected show invalid Reset sort

Date	Name	Res	Weight Avg	Adiron MP: 5.7 nd: 290 im0 im1 GT nonocc	ArtL MP: 1.5 nd: 256 im0 im1 GT nonocc	Jadepl MP: 5.2 nd: 640 im0 im1 GT nonocc	Motor MP: 5.9 nd: 280 im0 im1 GT nonocc	MotorE MP: 5.9 nd: 280 im0 im1 GT nonocc	Piano MP: 5.4 nd: 260 im0 im1 GT nonocc	PianoL MP: 5.4 nd: 260 im0 im1 GT nonocc	Pipes MP: 5.7 nd: 300 im0 im1 GT nonocc	Playrm MP: 5.3 nd: 330 im0 im1 GT nonocc	Playt MP: 5 nd: 290 im0 im1 GT nonocc	PlaytP MP: 5 nd: 290 im0 im1 GT nonocc	Recyc MP: 5.6 nd: 260 im0 im1 GT nonocc	Shelvs MP: 5.9 nd: 240 im0 im1 GT nonocc	Teddy MP: 2.7 nd: 256 im0 im1 GT nonocc	Vintge MP: 5.5 nd: 760 im0 im1 GT nonocc
09/28/15	R-NCC	F	**2.91** 1	**0.98** 1	**3.85** 1	**6.73** 1	**4.33** 1	**5.11** 1	**1.24** 1	**1.68** 1	**5.46** 1	**2.40** 1	**1.69** 1	**1.98** 1	**1.14** 1	**2.64** 1	**0.84** 1	**0.93** 1
10/07/14	IDR	H	8.41 2	5.85 5	6.71 3	18.8 3	5.82 2	6.00 2	5.48 3	4.99 2	12.5 4	5.78 2	37.6 29	4.51 2	4.21 3	7.59 3	2.99 2	8.56 5
03/15/16	MPSV	Q	8.70 3	6.39 7	5.76 2	17.7 2	6.13 3	6.24 3	8.38 15	9.79 9	10.4 2	6.68 4	27.7 19	4.79 3	5.50 9	12.0 11	5.15 8	8.70 6
12/18/15	INTS	H	9.09 4	5.47 2	8.56 6	29.6 5	8.10 5	8.89 5	5.08 2	6.10 3	13.3 5	6.27 3	10.6 4	5.56 5	3.63 2	10.4 6	3.82 4	9.95 9
07/28/14	SGM	F	9.84 5	10.6 16	6.89 4	24.1 4	7.34 4	7.44 4	6.31 7	7.87 6	10.8 3	8.50 5	38.4 30	6.02 7	4.84 6	7.00 2	4.13 5	7.18 2
11/12/14	LCU	Q	11.1 6	5.59 3	10.6 9	32.0 7	12.5 11	12.8 11	6.24 6	7.52 5	18.3 12	9.45 8	9.59 3	6.39 8	5.16 7	14.7 16	3.05 3	11.1 11
07/28/14	SGM	H	11.8 7	8.43 12	8.92 7	32.9 8	9.76 6	9.73 6	6.38 8	8.00 7	14.1 6	8.85 7	44.5 38	7.34 10	5.30 8	9.59 5	5.19 9	7.70 3
06/09/16	MPSVe1	Q	12.9 8	6.18 6	10.8 10	31.0 6	11.3 8	10.5 7	9.00 18	10.1 10	15.6 9	9.64 9	22.5 12	5.68 6	6.73 13	22.1 33	9.83 20	23.8 22

Full results (compared to 44 methods)

MPSVe1 = Multi-scale diffusion + Filtering

Training dense ; bad 4.0 ; all pixels

10/31/15	SPS	F	22.0 26	13.2 27	23.0 30	32.0 24	14.2 32	13.8 28	22.4 31	31.9 28	24.1 34	32.3 37	23.3 15	21.4 36	18.7 36	42.0 31	11.6 17	32.7 20
04/27/16	JEM	Q	22.3 27	9.60 21	26.7 35	37.6 33	15.4 35	16.4 33	20.7 26	34.6 30	20.1 19	30.1 29	26.6 20	22.0 37	13.0 26	38.4 22	14.6 34	35.0 25
06/09/16	MPSVe1	Q	22.6 28	14.8 30	21.2 25	31.8 22	14.1 31	12.6 25	21.3 28	29.7 25	25.0 36	29.9 28	47.8 38	16.5 27	15.0 31	46.7 39	13.4 29	38.3 32
07/28/14	Cens5	H	22.6 29	16.4 32	20.7 21	31.9 23	12.6 27	11.5 24	22.2 30	31.8 27	22.3 30	31.1 34	49.7 41	16.0 24	14.9 30	41.8 30	14.1 32	45.2 41
09/14/15	ELAS	H	23.6 30	14.8 29	20.8 24	42.3 37	14.5 33	14.0 29	23.1 33	37.3 32	21.3 23	31.2 35	44.3 34	15.6 22	19.0 38	42.8 34	14.7 35	35.2 27

Full results (compared to 44 methods)

MPSVe1 = Multi-scale diffusion + Filtering

Set: training dense | training sparse
Metric: bad 0.5 | bad 1.0 | bad 2.0 | bad 4.0 | avgerr | rms | A50 | A90 | A95 | A99 | time | time/MP | time/GD
Mask: nonocc | all

plot selected | show invalid | Reset sort

Date	Name (avgerr (pixels))	Res	Avg (Weight)	Adiron MP: 5.7 nd: 290 im0 im1 GT nonocc	ArtL MP: 1.5 nd: 256 im0 im1 GT nonocc	Jadepl MP: 5.2 nd: 640 im0 im1 GT nonocc	Motor MP: 5.9 nd: 280 im0 im1 GT nonocc	MotorE MP: 5.9 nd: 280 im0 im1 GT nonocc	Piano MP: 5.4 nd: 260 im0 im1 GT nonocc	PianoL MP: 5.4 nd: 260 im0 im1 GT nonocc	Pipes MP: 5.7 nd: 300 im0 im1 GT nonocc	Playrm MP: 5.3 nd: 330 im0 im1 GT nonocc	Playt MP: 5 nd: 290 im0 im1 GT nonocc	PlaytP MP: 5 nd: 290 im0 im1 GT nonocc	Recyc MP: 5.6 nd: 260 im0 im1 GT nonocc	Shelvs MP: 5.9 nd: 240 im0 im1 GT nonocc	Teddy MP: 2.7 nd: 256 im0 im1 GT nonocc	Vintge MP: 5.5 nd: 760 im0 im1 GT nonocc
05/12/16	PMSC	H	**4.90** 1	**1.02** 1	4.88 4	24.7 7	**2.98** 1	**2.94** 1	2.50 3	**4.39** 1	**6.27** 1	4.18 2	**1.81** 1	**2.00** 1	**1.34** 1	5.44 4	1.88 3	5.78 4
01/19/16	NTDE	H	5.23 2	1.38 2	4.74 3	25.8 9	3.26 2	3.35 2	**2.29** 1	4.60 2	7.17 2	**4.13** 1	3.63 2	2.68 3	1.58 2	5.06 3	2.16 5	**4.67** 1
11/12/14	LCU	Q	5.90 3	1.75 4	5.50 8	**19.7** 1	3.92 4	4.06 6	3.26 8	5.89 9	8.58 4	4.85 4	4.74 4	2.99 6	2.26 8	6.85 8	1.85 2	17.4 31
12/18/15	INTS	H	5.91 4	1.90 5	5.42 7	22.7 5	4.21 6	4.33 9	3.03 6	6.08 11	8.69 5	4.65 3	6.59 8	3.21 9	1.72 3	7.47 10	2.56 12	7.47 8
11/03/15	MC-CNN+RBS	H	6.67 5	2.22 7	8.42 23	22.2 4	3.95 5	3.87 3	2.34 2	4.74 4	13.9 32	9.76 25	4.80 5	3.66 15	2.38 9	4.63 2	5.90 30	5.13 2
04/12/16	MeshStereoExt	H	7.22 6	2.27 8	11.2 27	21.4 2	5.07 16	5.56 18	3.38 10	4.93 5	11.4 17	12.3 29	5.16 7	4.80 25	2.10 4	6.01 5	5.34 29	7.00 6
04/19/15	MeshStereo	H	7.59 7	2.39 9	6.44 13	36.4 24	5.40 22	5.71 20	3.25 7	5.45 7	11.6 18	6.34 11	4.92 6	2.73 4	2.25 7	11.1 28	1.90 4	5.62 3
07/28/14	SGM	H	7.63 8	2.49 13	5.04 5	25.9 10	4.38 8	4.15 7	3.56 12	6.62 12	9.47 8	5.33 5	26.8 39	3.40 11	2.84 14	8.82 14	2.30 7	16.1 27
05/28/16	APAP-Stereo	H	7.82 9	4.90 27	5.70 9	44.5 34	3.86 3	3.98 5	3.02 5	4.64 3	8.02 3	9.01 23	4.56 3	4.04 18	3.12 17	**3.36** 1	2.58 13	6.52 5
04/17/15	TMAP	H	7.88 10	2.44 11	5.88 10	30.9 14	4.72 11	4.41 10	3.86 14	13.7 29	9.25 7	6.12 8	16.6 26	3.13 7	2.22 6	10.9 26	2.73 15	10.5 14
11/05/15	GCSVR	H	8.19 11	2.13 6	9.75 26	33.3 20	4.40 9	4.58 12	2.87 4	6.07 10	10.7 12	5.86 7	13.4 18	3.19 8	2.21 5	11.1 29	**1.67** 1	18.8 38
04/08/15	REAF	H	8.49 12	4.07 21	7.29 19	32.4 17	4.76 12	4.70 13	5.00 22	11.7 27	10.9 13	8.61 20	15.1 22	4.04 18	2.54 11	9.72 21	3.30 21	9.29 13
07/25/14	SGM	Q	8.51 13	2.46 12	7.84 22	32.1 16	5.17 18	4.95 15	5.12 24	8.42 17	11.3 16	6.15 9	18.5 30	3.59 13	2.84 14	8.35 12	2.54 9	15.5 24
07/28/14	SGM	F	8.53 14	4.90 27	**4.65** 1	30.3 13	4.70 10	4.32 8	3.77 13	5.76 8	10.4 9	7.04 12	25.4 37	4.27 21	4.36 26	9.34 19	2.33 8	17.7 34
10/07/14	IDR	H	8.57 15	2.60 14	5.97 12	30.0 12	4.26 7	3.90 4	4.39 17	10.8 25	10.4 10	6.30 10	39.6 43	2.61 2	2.42 10	10.2 22	2.54 9	9.09 11
06/09/16	MPSVe1	Q	8.61 16	3.26 17	7.19 18	23.9 6	6.73 26	6.41 23	5.01 23	8.29 16	13.6 30	7.69 13	14.3 21	3.63 14	3.78 22	12.7 35	3.96 24	17.4 32

MIDDLEBURY 2014 BENCHMARK

Full results (compared to 44 methods)

MPSVe1 = Multi-scale diffusion + Filtering

Set: training dense | training sparse
Metric: bad 0.5 | bad 1.0 | bad 2.0 | bad 4.0 | avgerr | rms | A50 | A90 | A95 | A99 | time | time/MP | time/GD
Mask: nonocc | all

plot selected | show invalid | Reset sort

Date	Name (rms (pixels))	Res	Avg (Weight)	Adiron MP: 5.7 nd: 290 im0 im1 GT nonocc	ArtL MP: 1.5 nd: 256 im0 im1 GT nonocc	Jadepl MP: 5.2 nd: 640 im0 im1 GT nonocc	Motor MP: 5.9 nd: 280 im0 im1 GT nonocc	MotorE MP: 5.9 nd: 280 im0 im1 GT nonocc	Piano MP: 5.4 nd: 260 im0 im1 GT nonocc	PianoL MP: 5.4 nd: 260 im0 im1 GT nonocc	Pipes MP: 5.7 nd: 300 im0 im1 GT nonocc	Playrm MP: 5.3 nd: 330 im0 im1 GT nonocc	Playt MP: 5 nd: 290 im0 im1 GT nonocc	PlaytP MP: 5 nd: 290 im0 im1 GT nonocc	Recyc MP: 5.6 nd: 260 im0 im1 GT nonocc	Shelvs MP: 5.9 nd: 240 im0 im1 GT nonocc	Teddy MP: 2.7 nd: 256 im0 im1 GT nonocc	Vintge MP: 5.5 nd: 760 im0 im1 GT nonocc
11/12/14	LCU	Q	**16.8** 1	7.01 3	14.6 3	57.2 2	16.1 2	16.2 2	7.33 3	15.5 10	24.4 3	**12.9** 1	12.1 2	8.99 3	7.72 4	14.6 6	**5.12** 1	36.2 24
05/12/16	PMSC	H	17.4 2	**6.70** 1	19.1 17	72.6 9	**15.5** 1	**15.4** 1	8.49 8	13.7 6	**21.0** 1	16.0 5	**7.05** 1	8.48 2	**4.83** 1	12.9 4	8.39 13	25.4 11
01/19/16	NTDE	H	17.7 3	7.61 4	14.6 4	73.6 10	16.6 3	17.0 4	**6.08** 1	12.5 3	23.8 2	13.1 2	13.9 3	9.74 7	6.16 3	11.3 3	9.98 16	20.9 5
12/18/15	INTS	H	18.4 4	8.84 6	16.3 7	66.2 6	18.5 11	19.3 14	8.12 7	14.2 7	26.3 7	13.8 4	19.5 7	10.6 11	5.85 2	15.9 9	9.16 15	18.8 3
11/03/15	MC-CNN+RBS	H	20.9 5	10.1 8	21.1 23	**56.9** 1	16.8 4	16.7 3	6.49 2	13.3 4	39.1 36	36.0 27	18.8 5	13.9 20	9.55 13	11.2 2	21.8 30	17.1 2
07/28/14	SGM	H	21.2 6	10.1 10	14.5 2	68.3 7	17.8 6	17.5 8	10.5 13	16.2 11	25.9 6	16.6 6	55.9 40	11.0 15	11.0 18	18.6 17	7.85 9	34.1 19
04/19/15	MeshStereo	H	21.4 7	10.7 12	18.7 16	86.7 24	21.7 22	23.0 21	8.96 9	12.5 2	30.3 18	17.5 9	14.5 4	9.22 5	8.44 9	25.2 36	6.11 3	**16.8** 1
06/09/16	MPSVe1	Q	22.1 8	9.70 7	18.2 12	62.6 3	22.9 24	22.3 20	11.8 19	19.4 18	32.3 20	21.1 16	23.7 10	9.31 6	10.1 15	23.4 32	14.6 26	36.4 25

Main aspects of the proposed method

- ✔ Geometrical constraints such as low baseline, fronto-parallelism for segmented objects, and ordering constraint, are not required.
- ✔ We distinguish disparity measurement from disparity estimation.
- ✔ The method is quite robust to occlusions.
- ✔ The cost aggregation takes into account the segmentations of the left and right images of the stereo pair.
- ✔ The multi-scale extension improves the regional consistency of the disparity function, without affecting its precision.
- ✔ A morphological filter detects and prunes bad measures, by using the concept of disparity clusters and by analysing disparity distributions on a regional basis.
- ✔ Results are comparable to the state-of-the-art.
- ✔ The RMSE remains quite low compared to the ground truth, for both sparse and full results.
- ✔ Disparity maps are visually appealing.

Perspectives

- ➔ The algorithm will be optimised and applied to higher image resolutions. An increase in performance is expected.
- ➔ Disparity function should be better regularised across homogeneous regions.
- ➔ It would be interesting to ameliorate the multi-scale extension so as to take several hierarchies of segmentations into account during the aggregation phase, rather than relying on a single segmentation.

Thank you for your attention

References

1. Aydin, T., Akgul, Y.S.: Stereo depth estimation using synchronous optimization with segment based regularization. Pattern Recognition Letters **31**(15) (2010) 2389–2396
2. Prince, S.: Models for grids. In: Computer vision: models, learning, and inference. Cambridge University Press (2012)
3. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(11) (2001) 1222–1239
4. Sun, J., Li, Y., Kang, S.B., Shum, H.Y.: Symmetric stereo matching for occlusion handling. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005. Volume 2., IEEE (2005) 399–406
5. Yang, Q., Wang, L., Yang, R., Stewénius, H., Nistér, D.: Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling. IEEE Transactions on Pattern Analysis and Machine Intelligence **31**(3) (2009)
6. Facciolo, G., de Franchis, C., Meinhardt, E.: Mgm: A significantly more global matching for stereovision. In: Proceedings of the British Machine Vision Conference (BMVC), BMVA Press (2015) 90.1–90.12
7. Bobick, A.F., Intille, S.S.: Large occlusion stereo. International Journal of Computer Vision **33**(3) (1999) 181–200
8. Hirschmüller, H.: Stereo processing by semiglobal matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence (2008)
9. Zbontar, J., LeCun, Y.: Stereo matching by training a convolutional neural network to compare image patches. arXiv preprint 1510.05970 (2015)

10. Beucher, S., Meyer, F.: The morphological approach to segmentation: the watershed transformation. Optical Engineering **34** (1992) 433–481
11. Hosni, A., Bleyer, M., Gelautz, M.: Near real-time stereo with adaptive support weight approaches. In: International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT). (2010) 1–8
12. Cigla, C., Alatan, A.A.: Information permeability for stereo matching. Signal Processing: Image Communication **28**(9) (2013) 1072–1088
13. Vachier, C., Meyer, F.: The viscous watershed transform. Journal of Mathematical Imaging and Vision **22**(2-3) (2005) 251–267
14. Zabih, R., Woodfill, J.: Non-parametric local transforms for computing visual correspondence. In: Third European Conference on Computer Vision 1994 Proceedings, Volume II. Springer Berlin Heidelberg (1994) 151–158
15. Fua, P.: A parallel stereo algorithm that produces dense depth maps and preserves image features. Machine vision and applications **6**(1) (1993) 35–49
16. : Middlebury stereo database. http://vision.middlebury.edu/stereo/
17. Scharstein, D., Hirschmüller, H., Kitajima, Y., Krathwohl, G., Nešić, N., Wang, X., Westling, P.: High-resolution stereo datasets with subpixel-accurate ground truth. In: Pattern Recognition. Springer (2014) 31–42
18. Sinha, S.N., Scharstein, D., Szeliski, R.: Efficient high-resolution stereo matching using local plane sweeps. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE (2014) 1582–1589