



## **Overview:**

- What constitutes a good test dataset?
- Which content and aspects have to be included so that we really test. Are positive test cases (The Good) enough? How about using border cases (The Ugly) and negative test cases (The Bad)?
- We create checklists of hazards that should be in your test data so that it can really test the robustness of CV algorithms!

# **Dataset Survey:**

> A thorough dataset survey for stereo vision test datasets gives a good overview of the historic progress that the community has made. We identified 28 datasets between 2002 and 2017:



- > A generic checklist "CV-HAZOP" was used as a starting point to generate a specialized hazard checklist for stereo vision
- Each hazard from the checklist was searched for in the well-known datasets Freiburg, HCI, KITTI, Middlebury and Sintel





# Analyzing Computer Vision Data - The Good, the Bad and the Ugly Oliver Zendel (AIT) Katrin Honauer (HCI) Markus Murschitz (AIT) Martin Humenberger (AIT) Gustavo Fernandez Dominguez (AIT)

	IVICATING	Er
L.s. / No / Number	No l.s.	Hi
L. s. / Part of / Position	Part of l.s. is visible	L.9
L. s. / Less / Beam prop.	Focused beam	Sc
Medium / Less / Transparency	Medium is optically thicker than expected	Fc
Object / Less / Complexity	Object is less complex than expected	Sc
Object / No / Reflectance	Obj. has no reflectance	W
Object / As well as / Reflectance	Obj. has both shiny and dull surface	O im
Objects / Spatial aper. / Reflectance	Refl. creates a chaotic pattern	La
Obs. / Faster / Position	Observer moves too fast	In
Obs. / No / PSF	No optical blurring	In
	s. / No / Number s. / Part of / Position s. / Less / Beam prop. Vedium / Less / Transparency Object / Less / Complexity Object / No / Reflectance Object / As well as / Reflectance Objects / Spatial aper. / Reflectance Obs. / Faster / Position Obs. / No / PSF	s. / No / NumberNo I.ss. / Part of / PositionPart of I.s. is visibles. / Less / Beam prop.Focused beamMedium / Less / TransparencyMedium is optically thicker than expectedObject / Less / ComplexityObject is less complex than expectedObject / No / ReflectanceObj. has no reflectanceObject / As well as / ReflectanceObj. has both shiny and dull surfaceObjects / Spatial aper. / ReflectanceRefl. creates a chaotic pattern Obs. / Faster / PositionObs. / No / PSFNo optical blurring

Examples for identified hazards in the test datasets. Images taken from: KITTI 2015 [15], KITTI 2012 [7], Sintel [2], Sintel [2], Middlebury [22], Middlebury [22], HCI [10], HCI [10], Freiburg CNN [17] and Freiburg CNN [17]

seq_ambush_2_frame_0004
seq_ambush_2_frame_0012
seq_ambush_2_frame_0014
seq_ambush_4_frame_0011
seq_bamboo_2_frame_0011
seq_bandage_2_frame_0011
seq_market_5_frame_0002
seq_market_6_frame_0004
seq_mountain_1_frame_0031
seq_mountain_1_frame_0050
seq_shaman_3_frame_0001
seq_shaman_3_frame_0032
seq_sleeping_1_frame_0050

	Left	GT	SAD	SGBM	Elas	SPSS
			NO STATE			
			V- Con Con de			
			VI I			
		4			4.	4
1						
0						
			En State			

Evaluation Results for Sintel. Intensity images were taken from the original test dataset [2]

Images taken from the original test datasets (I.t.r, t.t.b): Middlebury [23], Middlebury [24], Neilson [16], EISATS S6 [20], New College [26], Pittsburgh [3], EVD [4], Ford Campus [18], HCI-Robust [11], KITTI 2012 [7], Leuven [12], Tsukuba [14], HCI-Synth [8], Stixel [19], Daimler Urban [25], Malaga Urban [1], Middlebury [22], Cityscapes [5], KITTI 2015 [15], MPI Sintel [2], Freiburg CNN [17], HCI Training [10], SYNTHIA [21], Virtual KITTI [6], Oxford RobotCar [13]

Contact: <u>oliver.zendel@ait.ac.at</u> Webpage: <u>www.vitro-testing.com</u>

### ighly underexposed image; only black-level noise in image is cut apart by image border cene with half lit object leaving a large portion severely underexposed g or haze in image reduces visibility depending on distance from observer cene contains simple object without texture or self-shading (e.g. grey opaque sphere) Vell-lit scene contains a very dark object without texture nor shading bject has a large glare spot on its surface that obscures same areas in the left/right arge parts of the image show an irregular distorted mirror-like reflection mage has parts with clearly visible motion blur nage contains strong aliasing artifacts







# Missing tests:

datasets.

## **Border cases (the Bad)**

- $\succ$  310: Two different sized objects are positioned on the same epipolar line but their projected views are identical
- > 694: Scene contains a clear reflection of observer together with potential matching parts on the same epipolar
- $\succ$  758: Scene contains pronounced refraction rings (e.g. oil slick)

### Negative test cases (the Ugly)

- 245: Cloud of visible particles (e.g. pollen, small leaves) in the air is obscuring the whole scene
- > 916: One camera lens contains dust/dried mud that creates a partially defocused area in the image
- > 933: Images contain rolling shutter artifacts
- All entries and more results on:



Correlation between hazard frames and algorithm performance Using only difficult frames still results in meaningful evaluations Testing with only easy frames will allow very little distinctions

> We identified over 60 hazard entries which were not found in any of the existing datasets and represent untested behavior. Of these we handpicked 16 border cases and 16 negative test cases which provide a good input to people who plan new challenging

# www.vitro-testing.com

