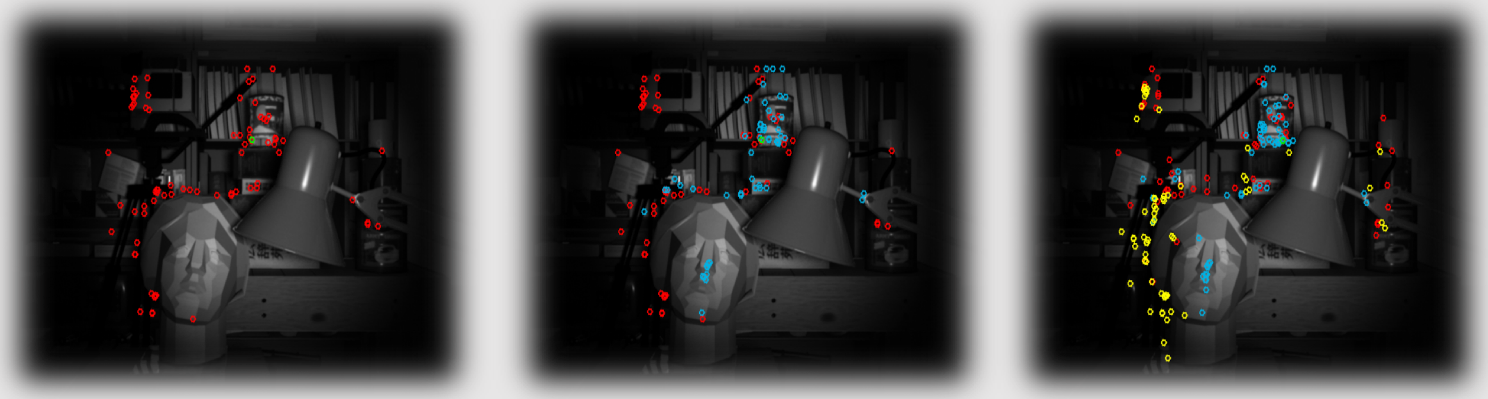
MULTIPLE LAYERS OF CONTRASTED IMAGES FOR ROBUST FEATURE-BASED VISUAL TRACKING



Xi WANG, Marc CHRISTIE, Eric MARCHAND

Univ Rennes, Inria, CNRS, IRISA

ICIP 2018 - ATHENS, GREECE

SLAM and Visual SLAM

SLAM = **S**imultaneous **L**ocalization **A**nd **M**apping

is the computational problem of constructing a map of an unknown environment while simultaneously keeping track of an agent's location within it

Autonomous Driving

(Image: Oxford Robotcar Dataset)

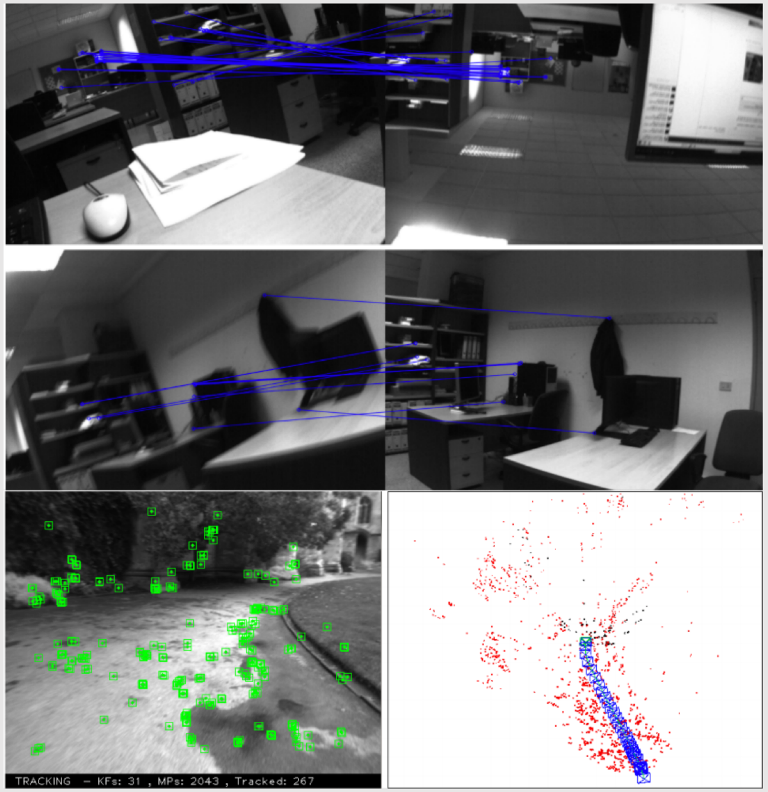
Indoor AR scene

(Image: AR Kit Apple)

SLAM and visual SLAM

Feature-based SLAM and low-level extractor matching

- Feature-based SLAMs rely on low-level extractors and descriptors
- A minimization of reprojection error of keypoints to estimate camera pose
- Insufficient extracting and wrong matching cause inaccurate estimations



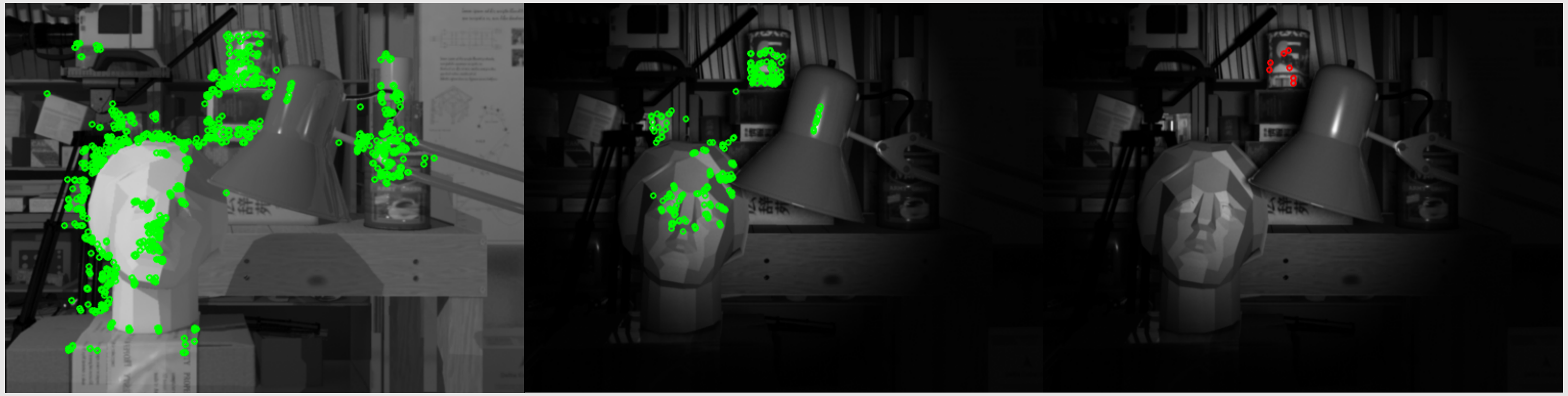
Feature-based ORB-SLAM [R. Mur-Artal et al. 2015]

Problematic: Illumination variance

BCA = **B**rightness **C**onstancy **A**ssumption

Daytime shift | **Season change** | **Bad exposure**

Feature detector : ORB from ORB-SLAM



(a) Daylight condition (b) Flashlight condition (c) Matching result

Synthetic Illumination Dataset: NewTsukuba Dataset [M. Peris-Martorell 2012]

Related Works

Illumination Invariant Imaging description

[M. Will et al. 2014] [R. Arandjelovic et al. 2016] [T. Nasser et al. 2018]

- Empirical or learning based illumination invariant description
- Incompatible to low-level feature extractors
- Remain high noise level

Automatic Exposure Correction

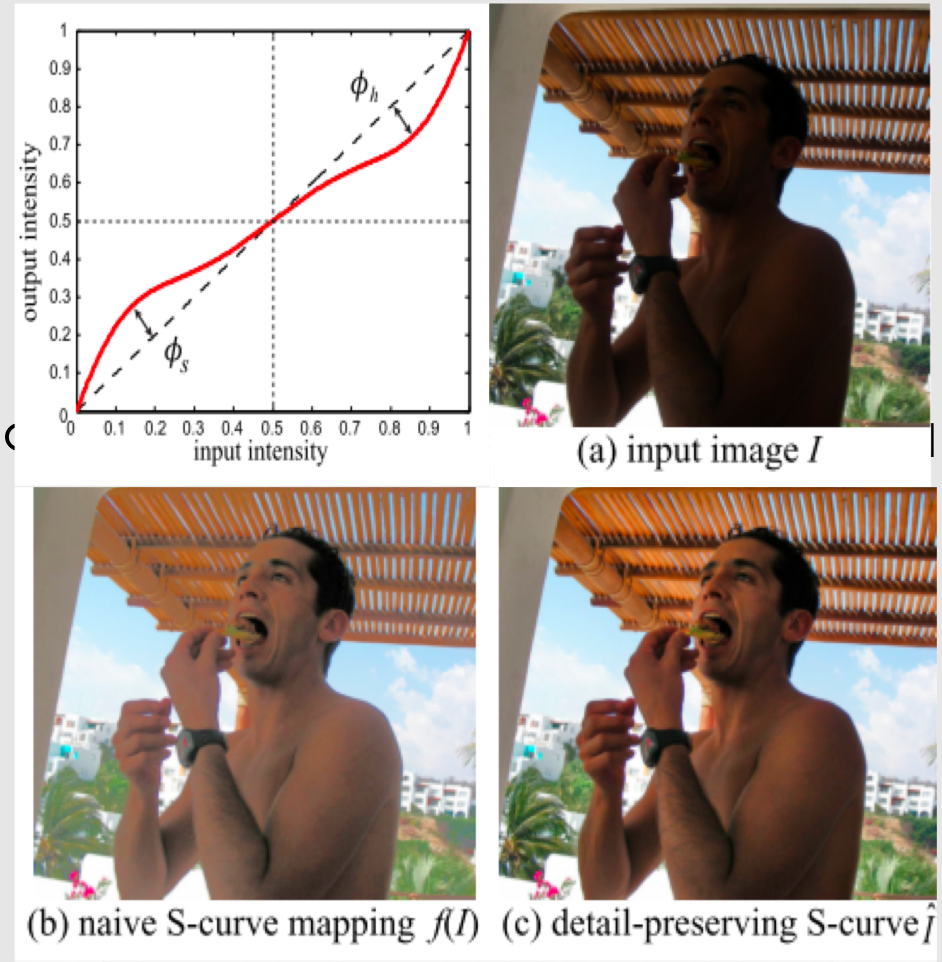
[Y. Lu et al. 2012] [S. Huang et al. 2013] [K. Gu et al. 2016]

- Learning a parameter of S-Curve or other correction function to enhance the contrast of whole image or segments of image
- Correction oriented towards aesthetic and human perception

HDR imaging

[T. Wang et al. 2011]

- Synthetic HDR from single image by changing parameters S Curve function
- Need specific designed computer vision tools for HDR image, time consuming



Automatic exposure correction with S-Curve function [Y. Lu et al. 2012]

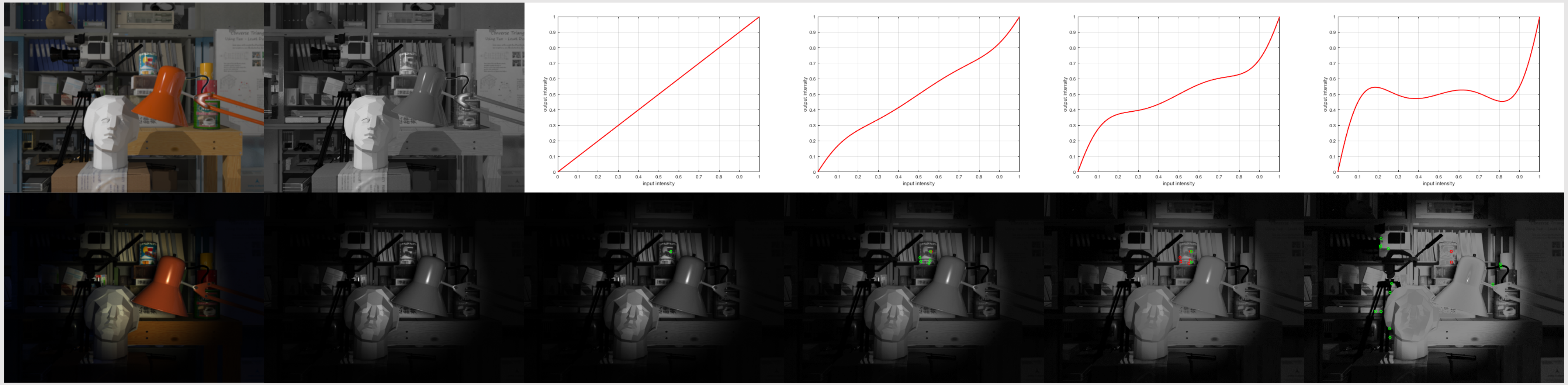
Illumination invariant image [M. Will et al. 2014]

No-Free-Lunch in contrast enhancement

S-Curve, a good idea?

Reference Image

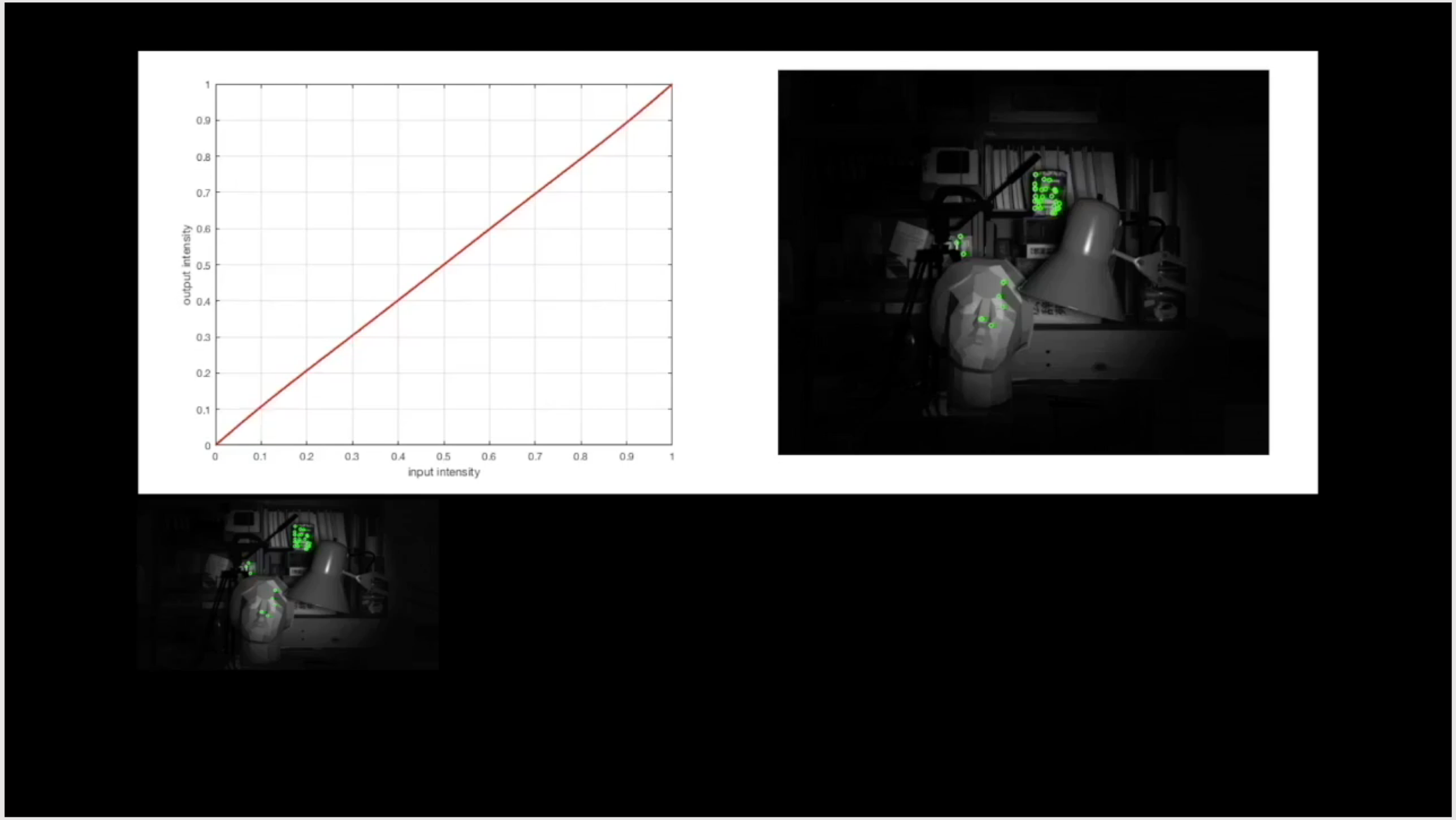
S-Curve Enhancement under different parameters



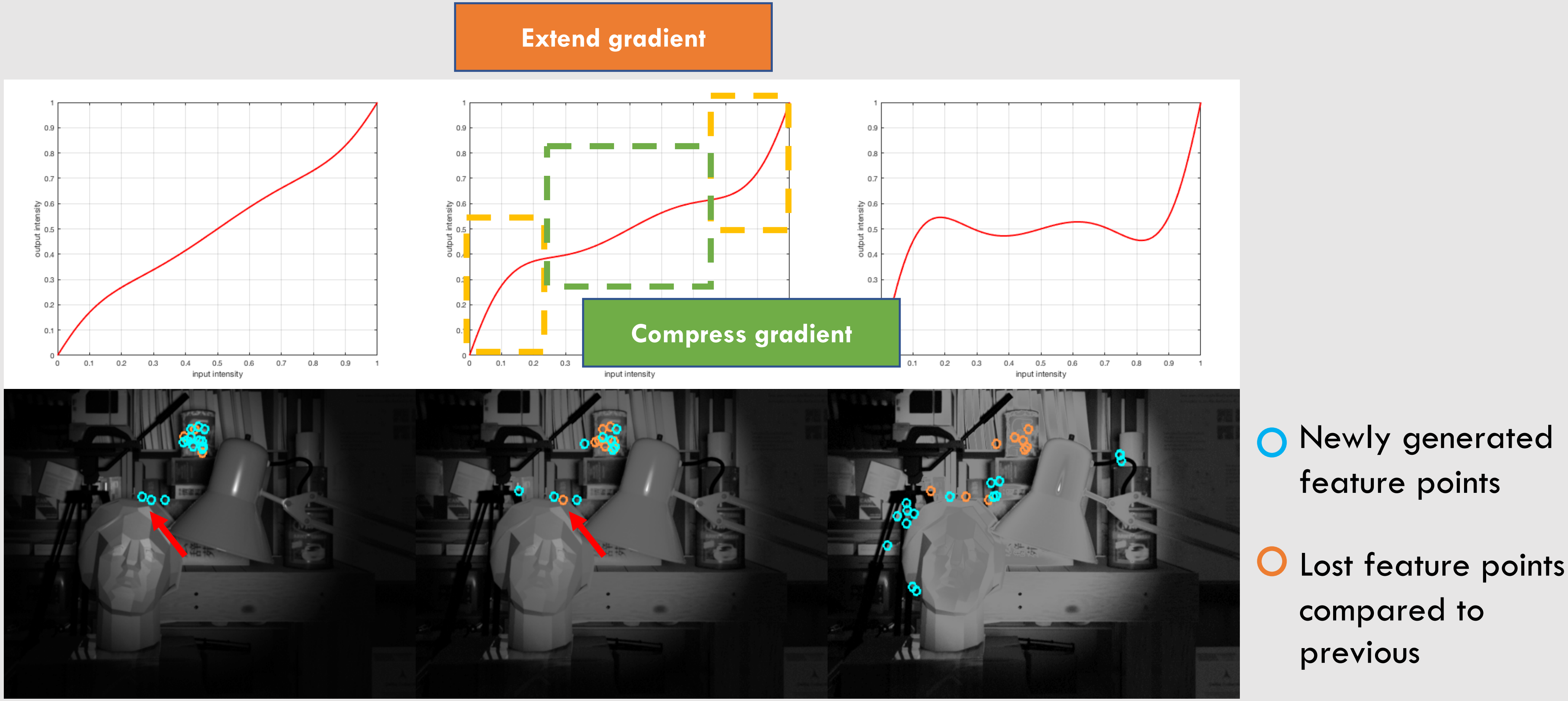
Query Image

No-Free-Lunch in contrast enhancement

S-Curve, a good idea?



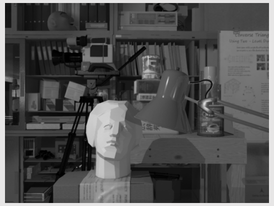
No-Free-Lunch in contrast enhancement



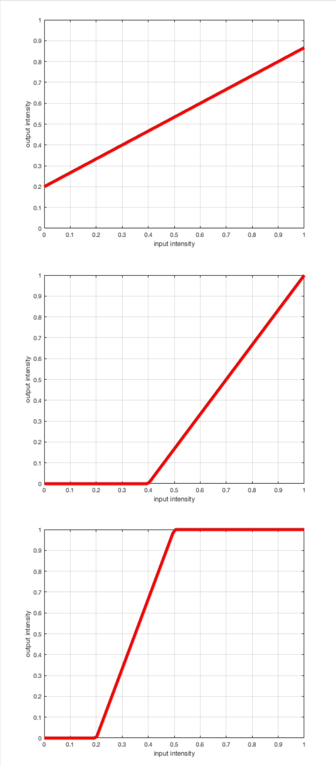
Feature points keep generating and losing across different parameters of S-Curve, which suggests a low-correlated information between each other.

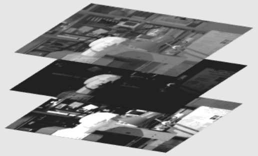
Our Approach: Multiple Layered Image (MLI)

Low-correlated contrasted layers

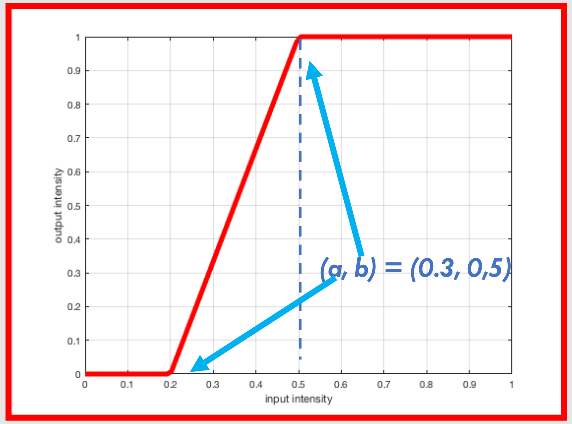


Searching for the best low-correlated contrast band parameters of SAT function **(a, b)**





MLI

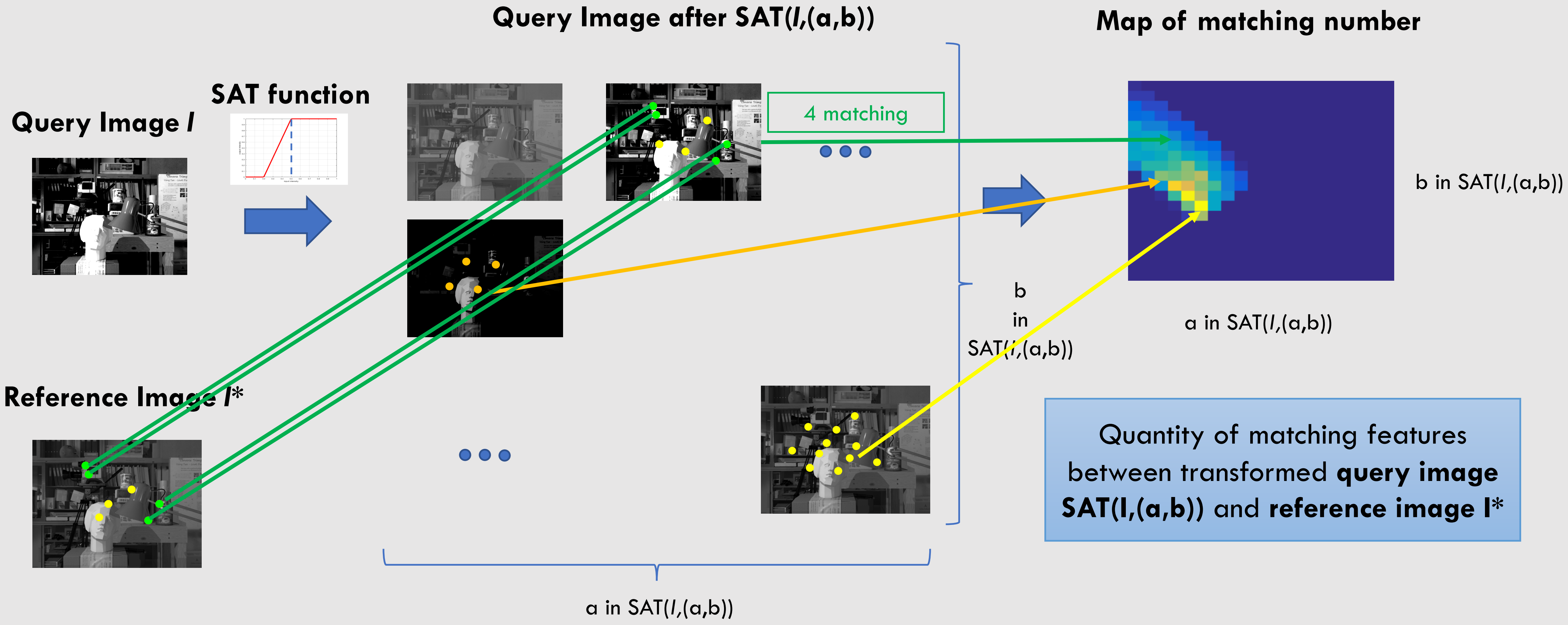


A saturated affine transform function **SAT(u=(a,b))** with contrast band (cutting points) of **SAT (0.3,0.5)**

Multiple layered image in which each layer represents the low-correlated information from specific contrast parameter.

An optimization frame

Generation Map of Matching Number

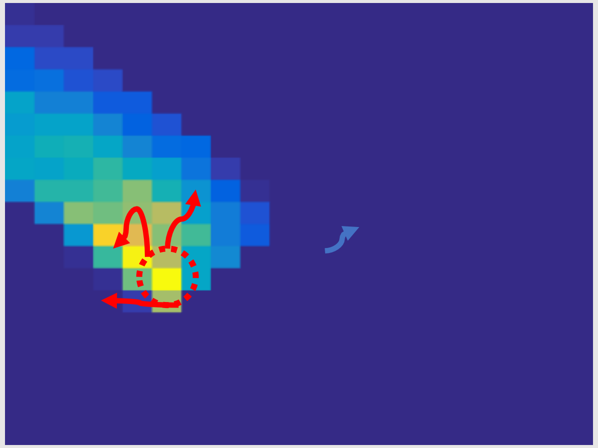


Quantity of matching features between transformed **query image SAT(I,(a,b))** and **reference image I***

An optimization frame

Generation Map of Similarity

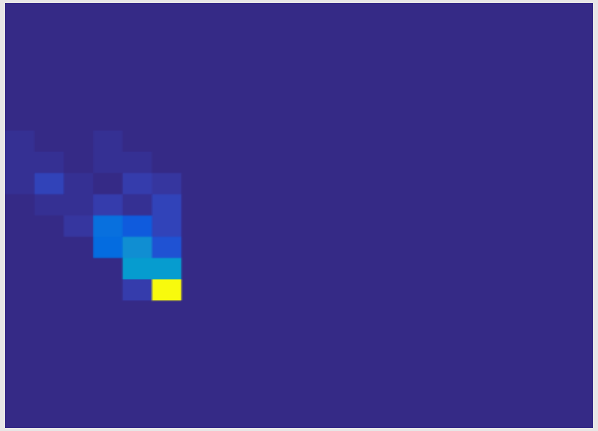
Map of matching number



We find **maximum** contrast band (a^*,b^*) as the configuration of **first** layer

Calculate correlation map of **maximum** against all others, yields this very contrast band's **<u>similarity map</u>**

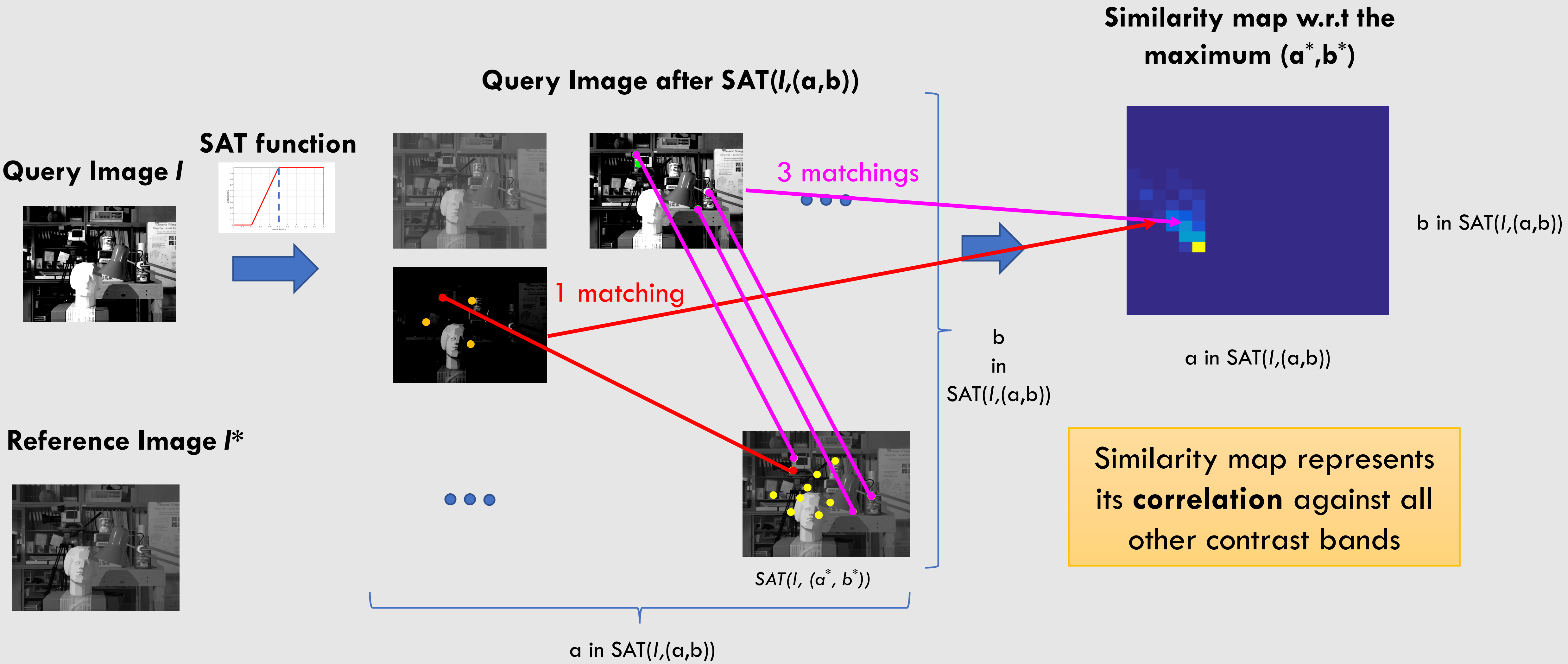
Similarity map w.r.t the maximum (a^*,b^*)



Similarity map represents one contrast band's **correlation** against contrast bands

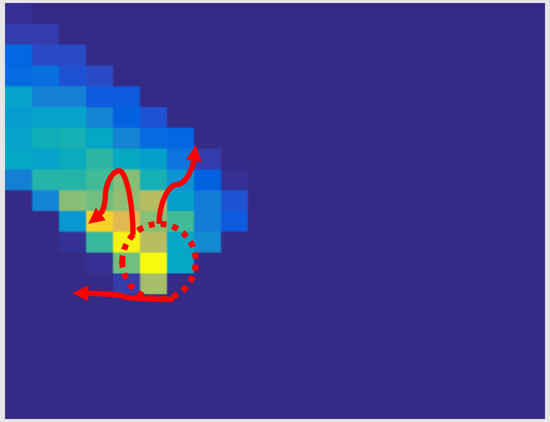
An optimization frame

Generation Map of Matching Number

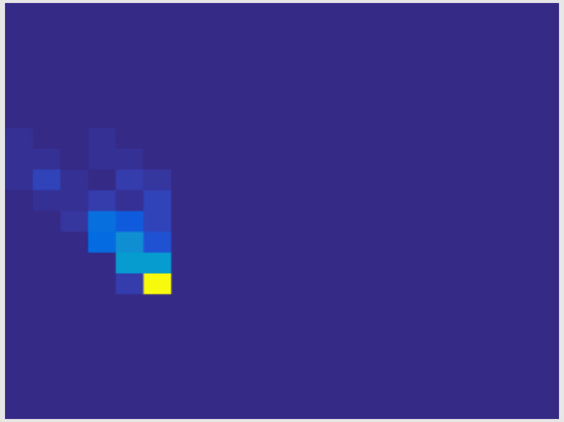


A matching map with low-correlation

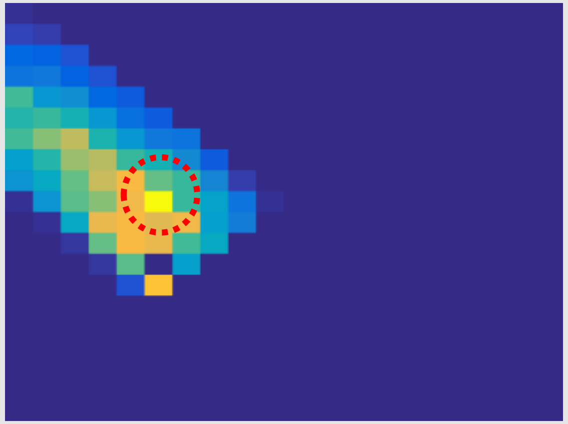
Map of matching number
Cost Function



Similarity represents its correlation with contrast bands



Map of matching number after subtraction of similarity
<u>Cost Function for next layer</u>



We find **maximum** contrast band (a,b) as the configuration of **first** layer

Iteratively generating contrast band for each layer

An iterative optimization framework

Algorithm 1 Optimal MLI

$i \leftarrow 0;\ C^0(\mathbf{u}) \leftarrow M_{Cor}(\mathbf{u});$

while $i = 0$ **or** $C^i(\mathbf{u}_i) > k * C^{i-1}(\mathbf{u}_{i-1})$ **do**

  $\mathbf{u}_i \leftarrow \mathrm{argmax}_{\mathbf{u}}(C^i(\mathbf{u}))$

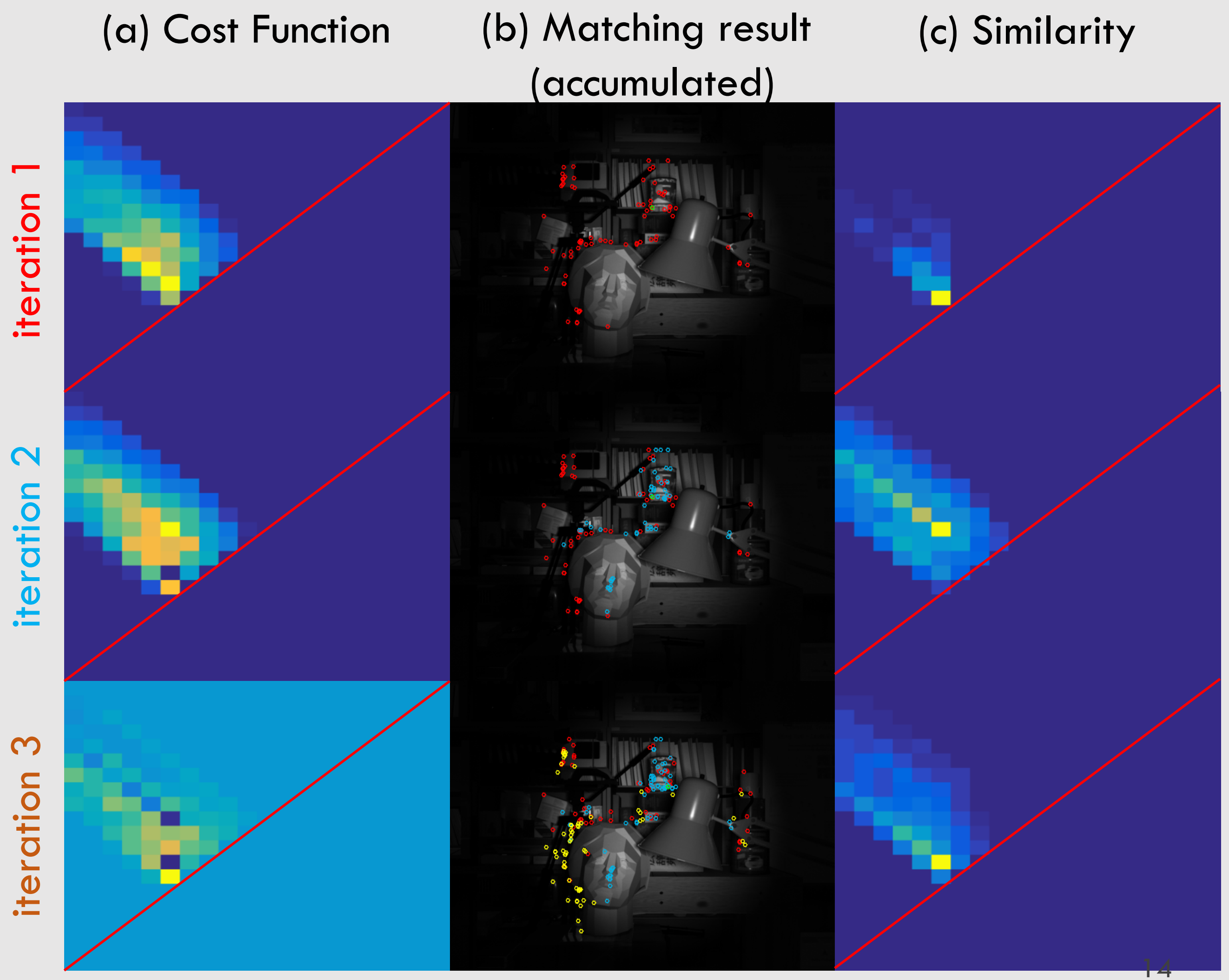
  $C^{i+1}(\mathbf{u}) \leftarrow C^i(\mathbf{u}) - M_{sim}^{\mathbf{u}_i}(\mathbf{u})$

  $i \leftarrow i + 1$

end while

return $\{\mathbf{u}_k\}_{k=1..N}$

Algorithm to calculate MLI



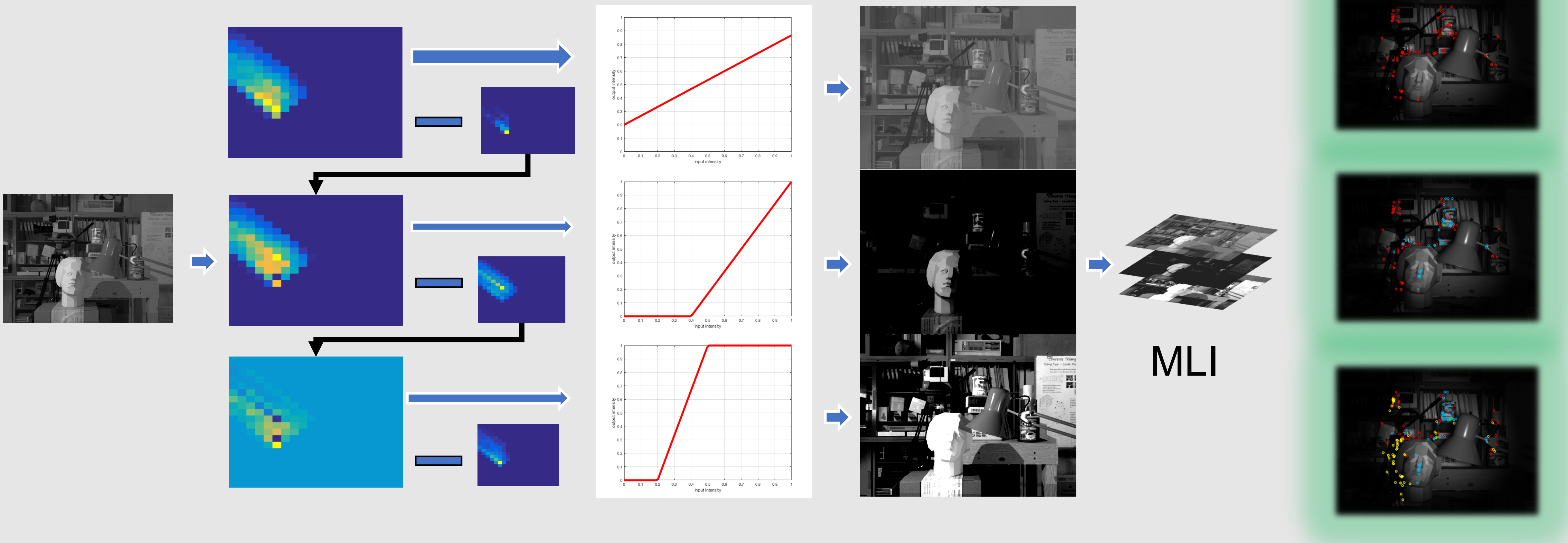
Keypoints from layer 1

Keypoints from layer 2

Keypoints from layer 3

Multiple Layered Image (MLI)

Pipeline with optimization framework



Multiple layered image in which each layer represents the low-correlated information from specific contrast parameter.

Results: Multiple Layered Image (MLI)

Experiments and results

Experiments

- Under four different lighting conditions of NewTsukuba synthetic data set [M. Peris-Martorell et al. 2012]
- Relocalization test: Keyframe retrieval from already generated map
- Repeatability test: Across different illumination conditions

daylight fluorescent lamp flashlight

Map generation Relocalization

Results: Multiple Layered Image (MLI)

Experiments and results

- MLI-supported ORB-SLAM gains better results against original version of ORB-SLAM and NID-SLAM [G. Pascoe et al. 2017]

- Repeatability test with ORB [E. Rublee et al. 2011], SURF [H. Bay 2006] and SIFT [D. Lowe 2004] extractors

V_2 \ V_1	Daylight			Fluo			Lamps			Flash		
	NID	ORB	MLI	NID	ORB	MLI	NID	ORB	MLI	NID	ORB	MLI
Daylight	99.3	100	**100**	96.7	96.2	**98.4**	73.9	**97.6**	53.6	74.6	**79.8**	77.1
Fluo	95.0	88.1	**95.1**	99.7	100	**100**	85.3	93.9	**100**	95.8	100	**100**
Lamps	88.3	55.7	**93.3**	**93.6**	79.8	93.4	93.1	100	**100**	84.3	37.9	**96.8**
Flash	23.8	30.7	**77.6**	92.2	90.6	**93.6**	0.00	0.00	**94.2**	92.0	**100**	99.3

cam \ ref		Daylight			Fluorescent			Lamps			Flashlight		
		ORB	SIFT	SURF	ORB	SIFT	SURF	ORB	SIFT	SURF	ORB	SIFT	SURF
Daylight	D	100/100	100/100	100/100	63.6/21.2	36.2/21.5	50.1/20.0	21.1/0.8	22.2/0.5	28.6/1.1	52.3/5.8	43.0/11.3	48.6/5.9
	M	100/100	100/100	100/100	**85.3/38.7**	**64.1/37.3**	**75.4/34.4**	**35.7/1.6**	**39.8/1.1**	**49.8/2.2**	**67.6/8.4**	**50.9/13.8**	**56.7/8.4**
Fluorescent	D	63.7/21.2	48.7/27.2	63.2/22.8	100/100	100/100	100/100	7.0/0.3	33.3/1.0	44.4/1.5	49.8/9.1	54.5/16.9	59.7/8.4
	M	**72.3/34.7**	**61.2/34.7**	**76.6/30.7**	100/100	100/100	100/100	**13.8/0.4**	**51.0/1.6**	**65.7/2.0**	**66.2/13.9**	**60.8/21.5**	**65.4/13.4**
Lamps	D	4.2/0.9	2.0/1.2	3.1/1.7	1.4/0.5	2.1/1.4	3.6/1.7	100/100	100/100	100/100	4.4/1.0	1.2/0.5	3.8/0.8
	M	**64.6/20.0**	**46.9/24.0**	**62.1/21.8**	**66.6/21.8**	**47.0/26.7**	**60.7/25.9**	100/100	100/100	100/100	**56.8/6.7**	**41.3/14.2**	**46.2/9.0**
Flashlight	D	34.0/5.3	12.2/5.4	16.4/5.3	32.4/7.6	11.3/6.3	16.2/6.2	14.6/0.5	5.1/0.0	12.1/0.2	100/100	100/100	100/100
	M	**58.1/11.8**	**31.4/11.8**	**44.6/11.4**	**61.4/16.6**	**30.5/15.0**	**44.0/13.5**	**18.4**/0.5	**16.4/0.5**	**35.0/0.8**	100/100	100/100	100/100

Table 1. Repeatability/matching ratio evaluation between MLI (M) and default single image (D) in percentage.

Conclusion and limitations

Contributions:

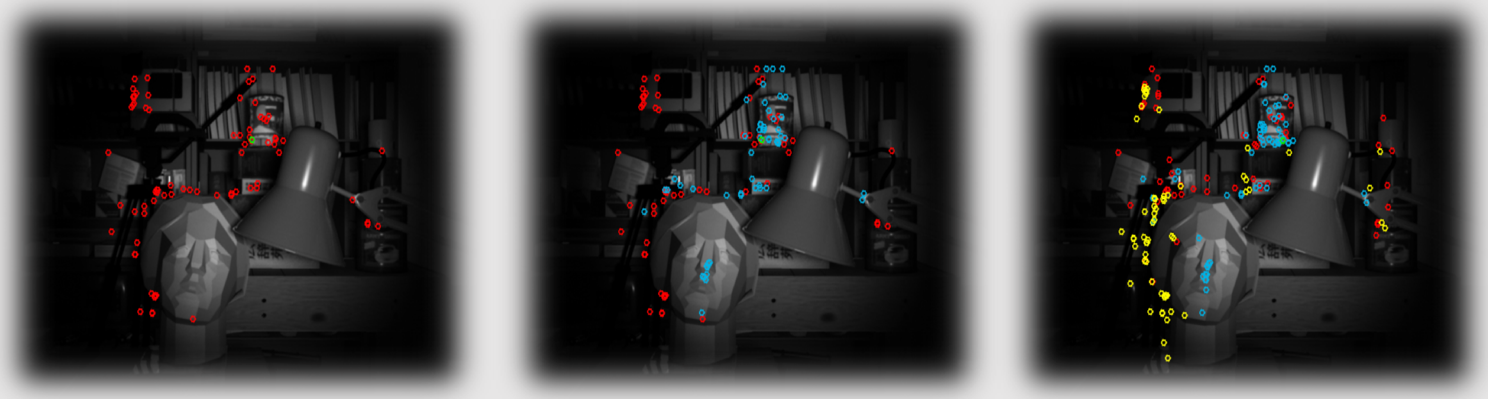
- Multiple layers of contrasted image (MLI) for more robust keypoint feature tracking
- A general solution for all extractors
- Quasi-real time SLAM implementation

Demonstration in video:

- ***Video***


M/Frame
M/Frame
ORB-SLAM
Proposed
Retracking the keyframes generated before
TRYING TO RELOCALIZE
KFs: 67 , MPs: 7853 , Tracked: 314
R:62 G:62 B:62
:108 G:108 B:108

MULTIPLE LAYERS OF CONTRASTED IMAGES FOR ROBUST FEATURE-BASED VISUAL TRACKING



Xi WANG, Marc CHRISTIE, Eric MARCHAND

Univ Rennes, Inria, CNRS, IRISA