How Does A Camera Look At One 3D CAD Object?

Chang Xing , **Chengjiang Long** , Hao Guo, Yongwei Nie , Yuan Zhang , Dehai Zhu , Qin Ma , Mengxiao Tian

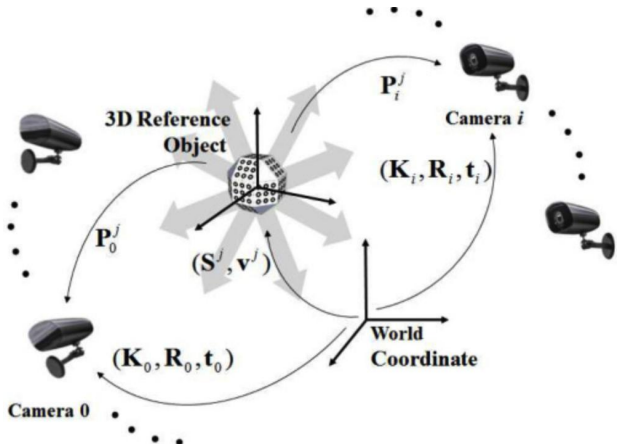
Moving robots

Flying drones

Q: How to guide the robots or drones to take pictures accurately?

A: To explore how does a camera looks at the objects.

Camera projection

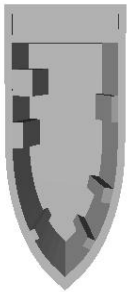


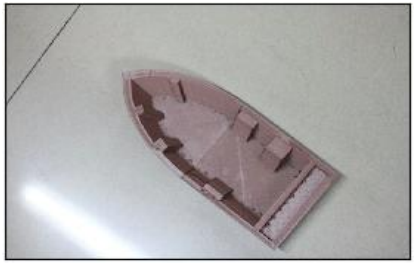
$$z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Internal calibration

External calibration

Problem definition





2D real image

?

$$\begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{11} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \end{bmatrix}$$

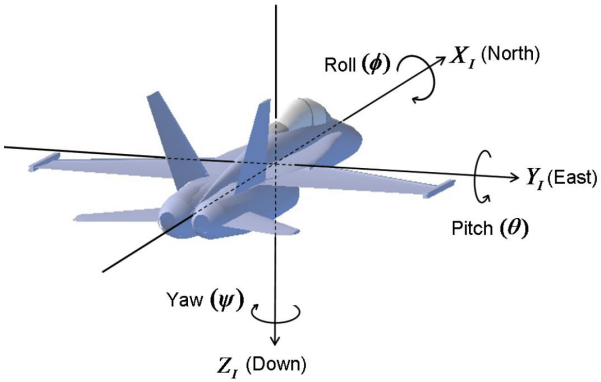
Predict RT

Generate synthetic image

6-DOF RT $(\theta_x, \theta_y, \theta_z, t_x, t_y, t_z)$

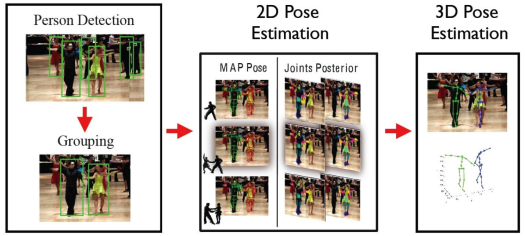
Rotation angles

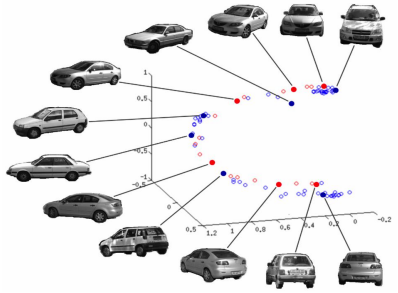
Translate vector



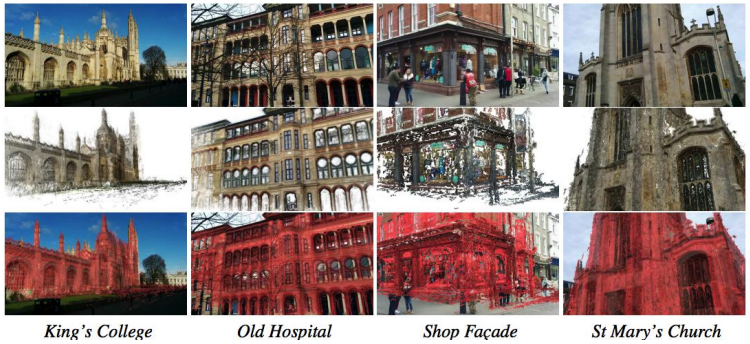
Related work

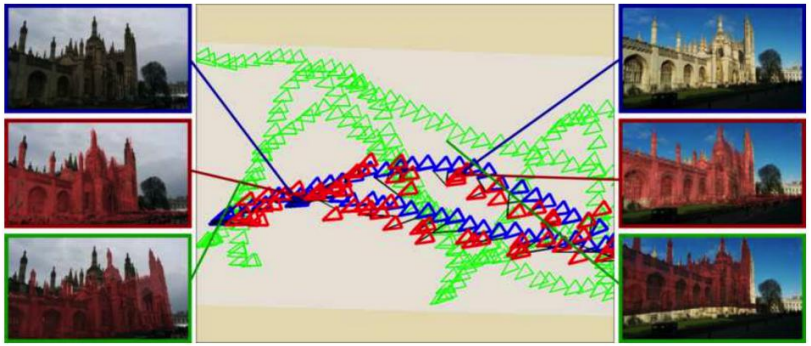
- Object pose estimation ([Pepik CVPR2012], [Teney CCRV2013] …)



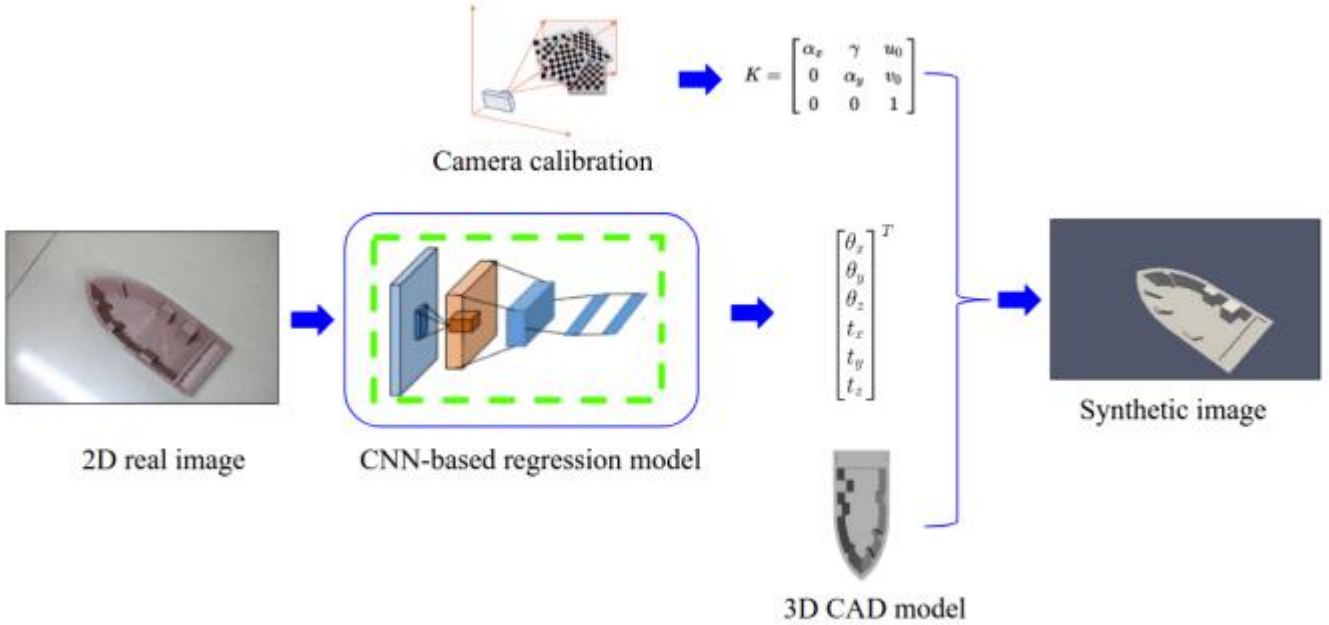


- PoseNet ([Kendll CVPR2015])

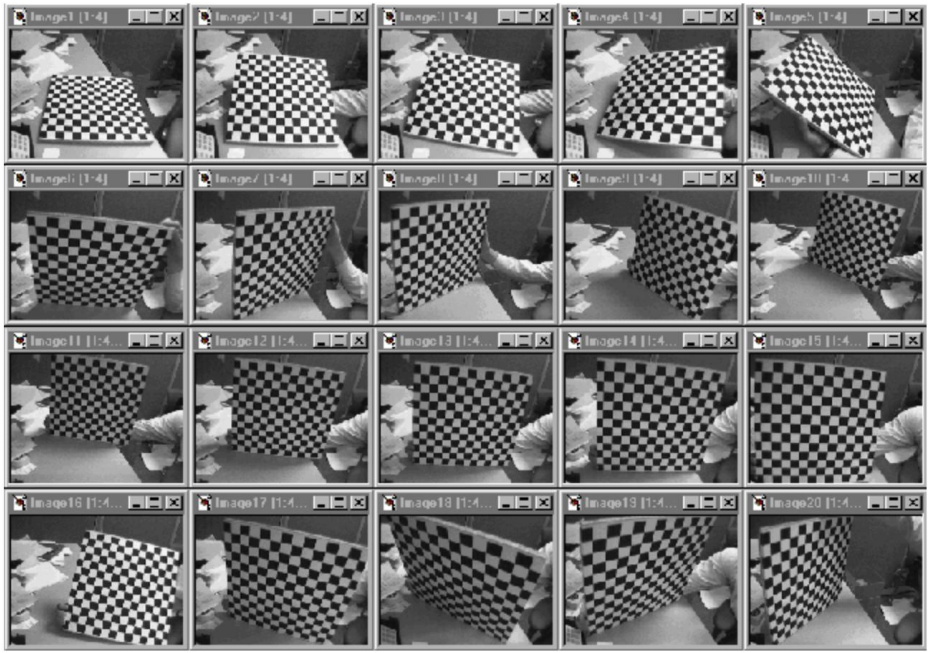




Pipeline


Camera calibration
$K = \begin{bmatrix} \alpha_x & \gamma & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$
2D real image
CNN-based regression model
$\begin{bmatrix} \theta_x \\ \theta_y \\ \theta_z \\ t_x \\ t_y \\ t_z \end{bmatrix}^T$
3D CAD model
Synthetic image

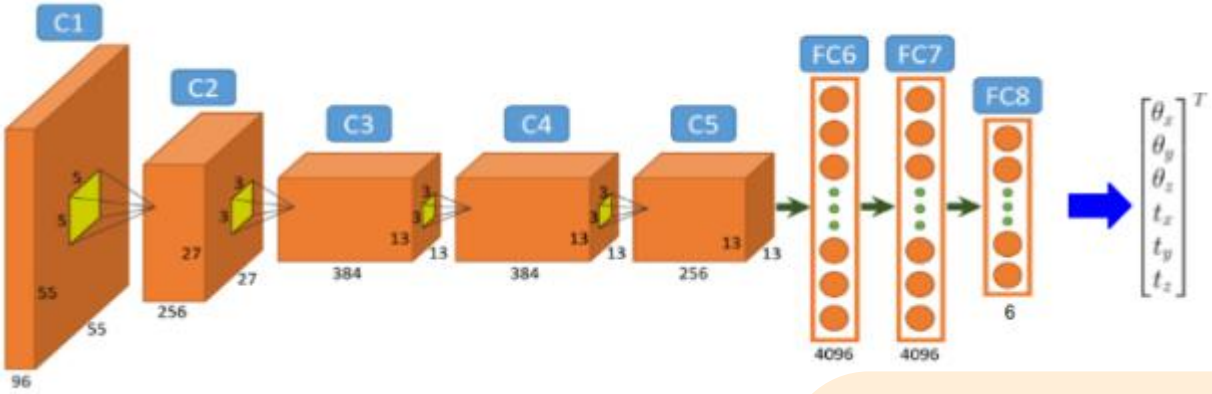
Camera calibration



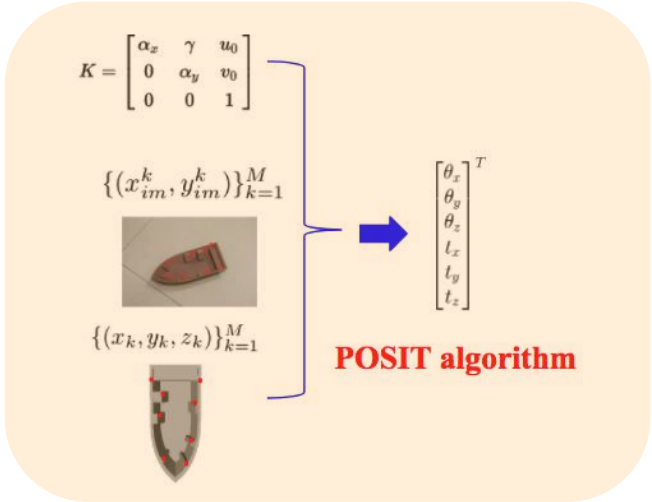
20-25 check board images.

Run a Camera Calibration Toolbox for Matlab.
http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example.html

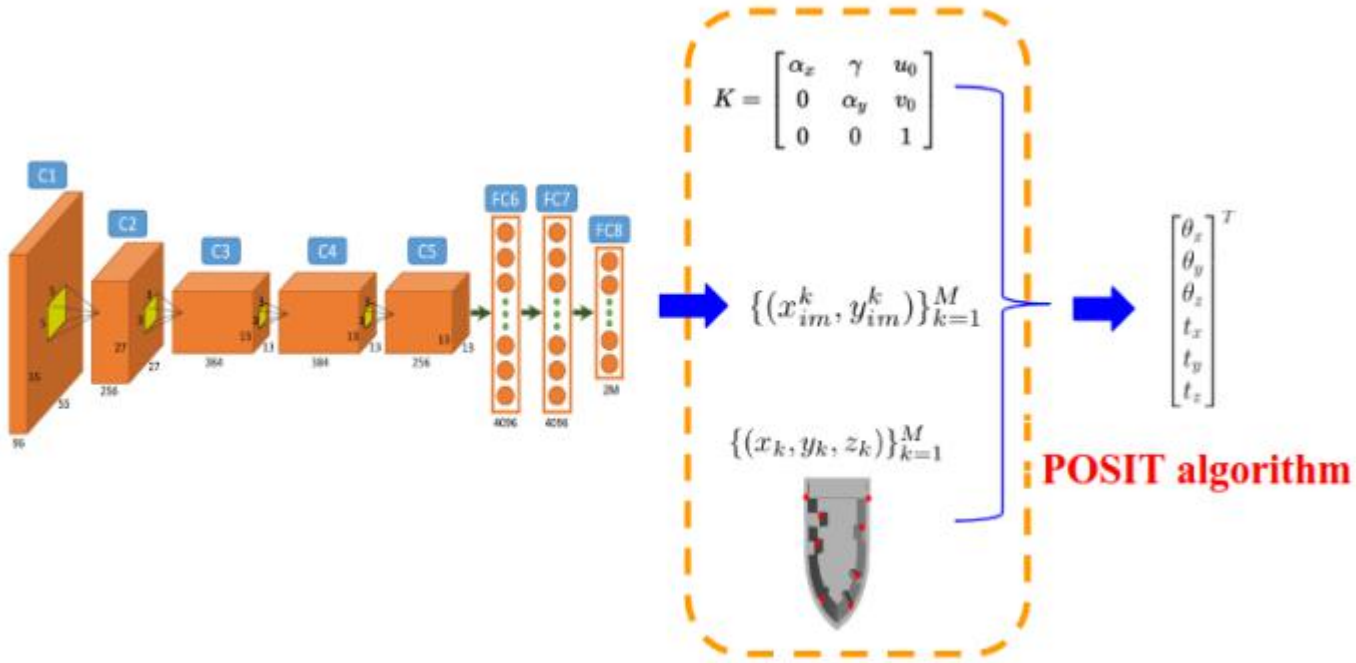
CNN-direct



- Indirect annotation: POSIT algorithm [Dementhon 1995]
- Modify the number of output channel to 6.
- Loss function: Euclidean loss.
- Issue between $-\pi$ and π.

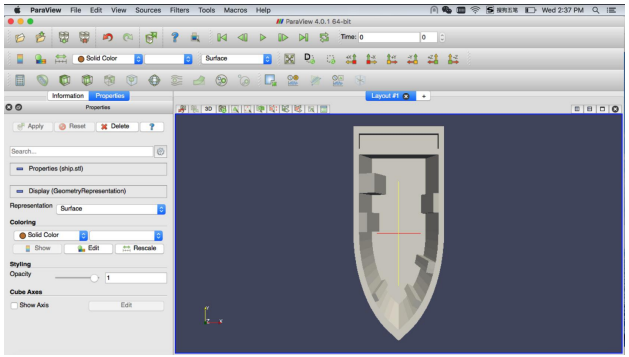


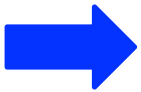
CNN-indirect

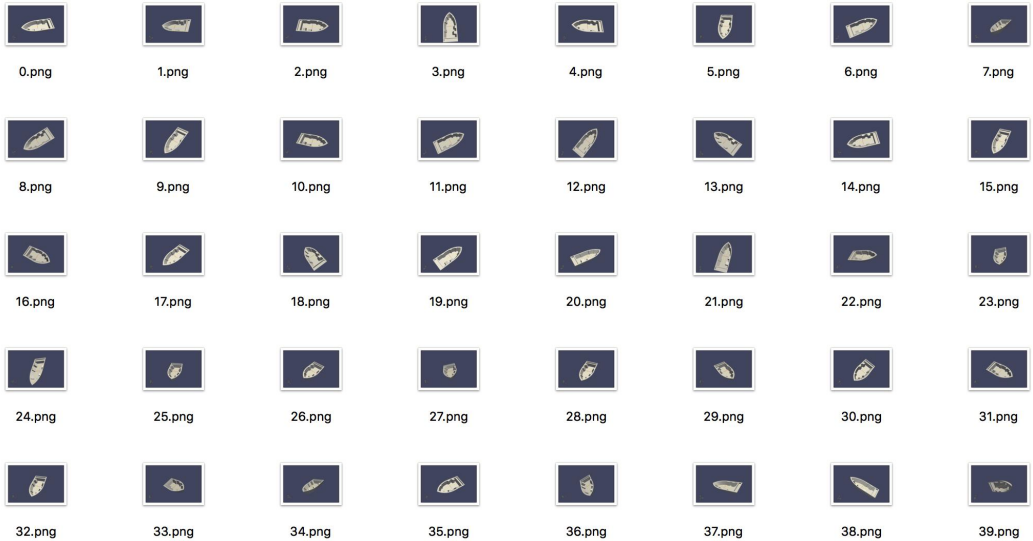


- Solve the Issue between $-\pi$ and π.

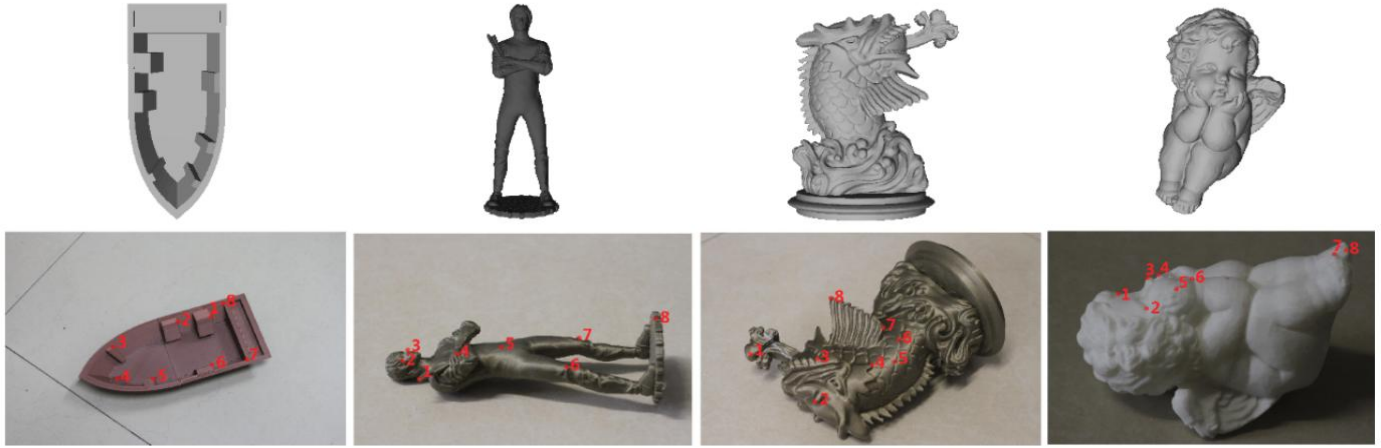
Visualization with Paraview







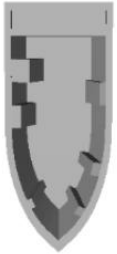
Experiment datasets

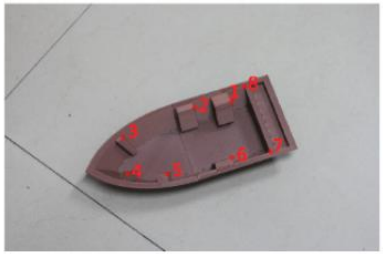


THE BRIEF INFORMATION FOR EACH CAD MODEL AND REAL IMAGES USED FOR EXPERIMENTS.

	small ship	Bruce Lee	fish statue	boy angel
Vertex number	9,378	703,248	233,184	554,082
Triangle number	3,126	234,416	77,728	184,694
Image nubmer	1,286	1,480	1,582	1,643
Training number	1,157	1,332	1,423	1,478
Testing number	129	148	159	165
θ_x range	[-3.137, 3.136]	[-3.141, 3.140]	[-3.141, 3.141]	[-3.104, 3.138]
θ_y range	[-1.128, 1.367]	[-1.239, 1.045]	[0.505, 1.557]	[-1.402, 0.512]
θ_z range	[-3.139, 3.141]	[-3.136, 3.131]	[-3.129, 3.124]	[-3.135, 3.132]
t_x range	[0.648, 11.828]	[0.358, 7.061]	[4.632, 222.456]	[19.033, 156.619]
t_y range	[0.417, 8.659]	[0.048, 3.955]	[0.637, 164.567]	[6.471, 131.387]
t_z range	[14.983, 39.193]	[2.642, 21.057]	[194.158, 776.526]	[216.857, 635.711]

Results on the ship model





THE REGRESSION ERRORS OF LANDMARKS ON THE SMALL SHIP MODEL.

	SVM-indirect		CNN-indirect	
Landmark	x	y	x	y
1	486.644	443.346	103.879	70.935
2	375.221	336.351	85.591	64.372
3	635.839	532.093	101.433	81.571
4	609.213	599.830	104.455	83.095
5	427.491	396.943	89.598	71.911
6	480.253	367.553	93.728	74.155
7	683.239	538.170	118.916	89.815
8	718.722	601.149	118.915	90.845
Average	552.078	476.929	102.064	78.337

$< 2\%$

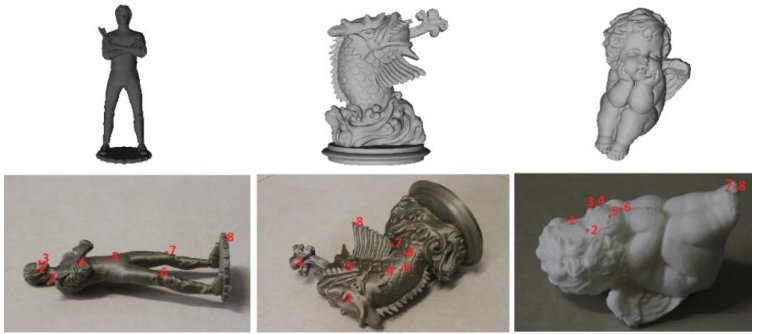
5184×3465

THE REGRESSION ERROR OF RTS ON THE SMALL SHIP MODEL.

	SVM		CNN	
Landmark	indirect	direct	indirect	direct
θ_x	2.270	2.102	0.866	**0.515**
θ_y	0.687	0.403	0.167	**0.128**
θ_z	1.659	1.425	**0.311**	0.368
θ_{avg}	1.539	1.310	0.448	**0.337**
t_x	3.427	1.209	**0.540**	0.561
t_y	2.257	1.010	0.469	**0.403**
t_z	15.633	2.951	3.652	**1.632**
t_{avg}	7.106	1.723	1.553	**0.865**

$19.3°$

Results on the other three models



THE REGRESSION ERROR OF RTS ON THE "BRUCE LEE" MODEL.

	SVM		CNN	
Landmark	indirect	direct	indirect	direct
θ_x	1.290	1.468	0.532	**0.407**
θ_y	0.297	0.221	0.089	**0.076**
θ_z	0.855	0.921	**0.132**	0.174
θ_{avg}	0.812	0.870	0.251	**0.219**
t_x	2.030	0.578	**0.188**	0.237
t_y	1.360	0.516	**0.122**	0.156
t_z	8.978	1.139	0.907	**0.713**
t_{avg}	4.123	0.744	0.406	**0.369**

12.5°

THE REGRESSION ERROR OF RTS ON THE FISH STATUE MODEL.

	SVM		CNN	
Landmark	indirect	direct	indirect	direct
θ_x	1.035	0.703	0.510	**0.209**
θ_y	0.311	0.146	0.073	**0.057**
θ_z	0.561	0.415	0.273	**0.143**
θ_{avg}	0.635	0.421	0.285	**0.136**
t_x	55.663	24.368	**5.916**	6.996
t_y	37.954	24.319	**3.867**	6.803
t_z	234.310	42.035	32.196	**16.690**
t_{avg}	140.514	30.241	13.993	**10.163**

7.8°

THE REGRESSION ERROR OF RTS ON THE BOY ANGEL MODEL.

	SVM		CNN	
Landmark	indirect	direct	indirect	direct
θ_x	0.475	0.300	0.186	**0.125**
θ_y	0.374	0.210	0.127	**0.095**
θ_z	0.922	0.910	0.234	**0.201**
θ_{avg}	0.590	0.473	0.182	**0.140**
t_x	46.828	16.157	6.514	**5.769**
t_y	30.575	15.234	**3.986**	4.969
t_z	205.063	35.164	34.995	**17.368**
t_{avg}	94.155	22.185	15.165	**9.369**

8.0°

Visualization on the ship model

Image | SVM-indirect | SVM-direct | CNN-indirect | CNN-direct | Ground-truth RT

Visualization on the “Bruce Lee” model

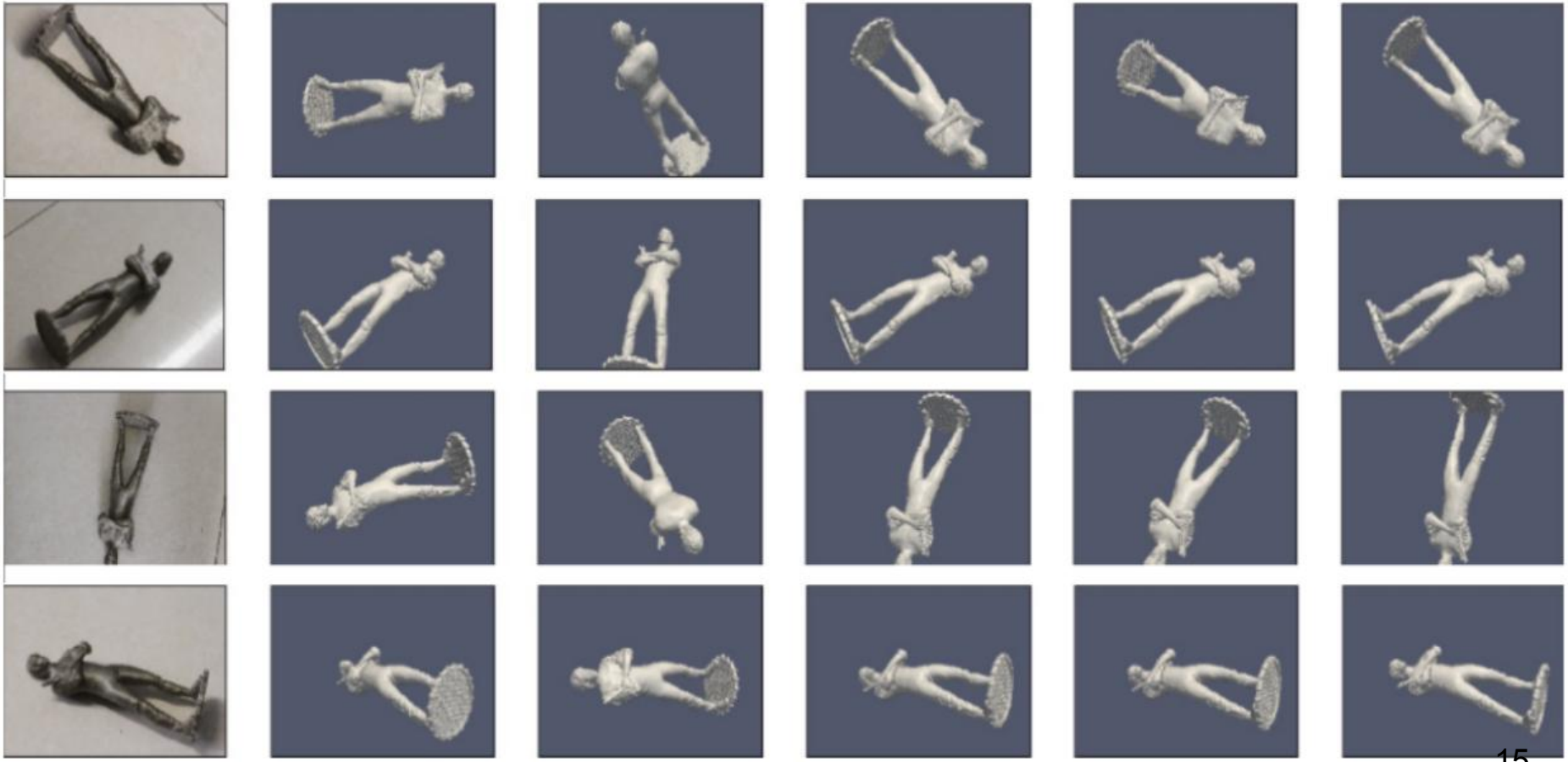


Image | SVM-indirect | SVM-direct | CNN-indirect | CNN-direct | Ground-truth RT

Visualization on the fish statue model

Image | SVM-indirect | SVM-direct | CNN-indirect | CNN-direct | Ground-truth RT

Visualization on the boy angel model

Image | SVM-indirect | SVM-direct | CNN-indirect | CNN-direct | Ground-truth RT

Angle ambiguous issue

Image CNN-indirect CNN-direct

Conclusion and future work

- We propose a CNN-based multi-output regression framework to estimate the camera's RTs directly and indirectly from images.
- Ware able to generate the synthetic images to visually verify the correctness and interpret how the camera looks at the 3D CAD object effectively and accurately.
- Our future work includes investigating and developing more powerful CNN regression models to reduce the regression errors, extending the current setting from a single CAD object to multiple CAD objects.

Thanks!

Email: chengjiang.long@kitware.com
Web: http://www.chengjianglong.com