Additional Material: Predicting Structured Geometry From Single And Multiple Views

Anonymous ECCV submission

Paper ID 1485

Abstract. We provide additional theoretical and experimental results that were omitted from the main paper due to space constraints.

1 Further Proof Details

Here we prove the equivalence of minimisation problems (20) and (29) stated in the main paper. For clarity we re–state the claim in the following proposition.

Proposition 1. *Let* $\boldsymbol{w}, \boldsymbol{\xi}$ *be the solution to*

$$\begin{aligned} \min_{\boldsymbol{w},\boldsymbol{\xi}} \;& \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{k=1}^{n}\xi_k \\ s.t. \;& \forall k, S \neq S_k : \; \big\langle \boldsymbol{w}, \Psi(\Theta_k, S_k)\big\rangle - \big\langle \boldsymbol{w}, \Psi(\Theta_k, S)\big\rangle \geq \Delta(S, S_k) - \xi_k \;. \end{aligned} \tag{1}$$

Let $\boldsymbol{w}', \boldsymbol{\xi}'$ *be the solution to*

$$\begin{aligned} \min_{\boldsymbol{w}',\boldsymbol{\xi}'} \;& \frac{1}{2}\|\boldsymbol{w}'\|^2 + \eta C\sum_{k=1}^{n}\xi'_k \\ s.t. \;& \forall k, S \neq S_k : \; \big\langle \boldsymbol{w}', \Psi(\Theta_k, S_k)\big\rangle - \big\langle \boldsymbol{w}', \Psi(\Theta_k, S)\big\rangle \geq \eta\Delta(S, S_k) - \xi'_k \;. \end{aligned} \tag{2}$$

Then

$$\boldsymbol{w}' = \eta\boldsymbol{w} \tag{3}$$

$$\boldsymbol{\xi}' = \eta\boldsymbol{\xi} \;. \tag{4}$$

Proof. Substituting for $\boldsymbol{w}'$ and $\boldsymbol{\xi}'$ in (2):

$$\begin{aligned} \min_{\boldsymbol{w},\boldsymbol{\xi}} \;& \frac{1}{2}\|\eta\boldsymbol{w}\|^2 + \eta C\sum_{k=1}^{n}\eta\xi_k \\ s.t. \;& \forall k, S \neq S_k : \; \big\langle \eta\boldsymbol{w}, \Psi(\Theta_k, S_k)\big\rangle - \big\langle \eta\boldsymbol{w}, \Psi(\Theta_k, S)\big\rangle \geq \eta\Delta(S, S_k) - \eta\xi_k \;. \end{aligned} \tag{5}$$

Dividing the objective by η^2 and the constraints by η we obtain the desired result. □

2 Further Experimental Details

In the main paper we described two loss functions: the relative depth error and the labelling error. We also described two feature spaces: one containing exclusively single view features and one containing and multiple view features. In total we trained four predictors, corresponding to all combinations of feature space and loss function. In this paper we provide further comparisons between these four predictors, and for clarity we will use the following notation throughout this document.

$f^{\mathsf{sview}}_{\mathsf{depth}}$	Trained w.r.t. Δ_{depth} in the single view feature space
$f^{\mathsf{sview}}_{\mathsf{labelling}}$	Trained w.r.t $\Delta_{\mathsf{labelling}}$ in the single view feature space
$f^{\mathsf{mview}}_{\mathsf{depth}}$	Trained w.r.t. Δ_{depth} in the multiple view feature space
$f^{\mathsf{mview}}_{\mathsf{labelling}}$	Trained w.r.t. $\Delta_{\mathsf{labelling}}$ in the multiple view feature space

Figure 1 shows the training evolution of the weight vector $\boldsymbol{w}$ for each of the four predictors. Due to the size of the single view feature space we show selected features only. The single view learning problem is, as expected, the more difficult problem, as shown by the considerably longer training times and the higher volatility during the exploration phase.

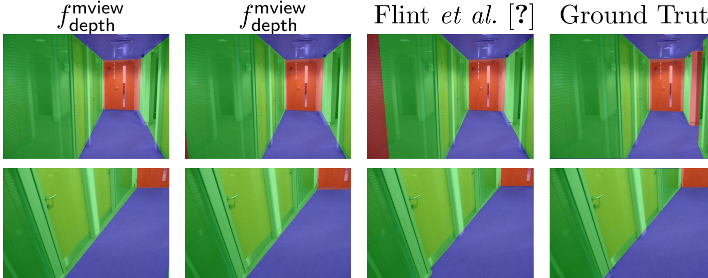
In the main paper we reported performance for each predictor on held–out training data. We reported only the performance metric corresponding to the loss for which each predictor was trained. Here we report performance for both metrics, for all predictors. That is, for each predictor, we compute both the depth and labelling error, regardless of which loss the predictor was trained against. These results are summarized in Figure 2.

Finally we show examples of predictions made by our system and the comparison systems in Tables 1 and 2.

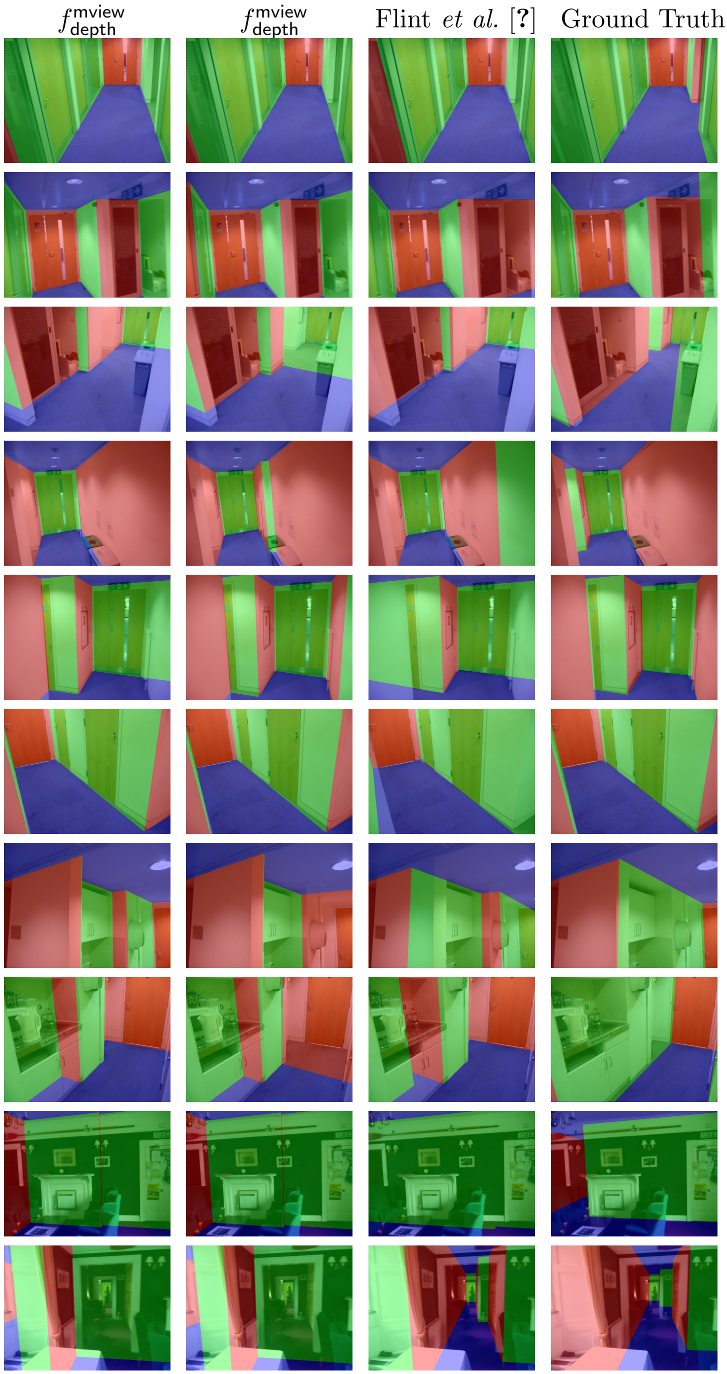
2.1 Examples of outputs from our system

Table 1: Reconstructions predicted by our system using the multiple view feature space; comparison with Flint *et al.* [?]. Instances in this figure are samples from the hold–out set.

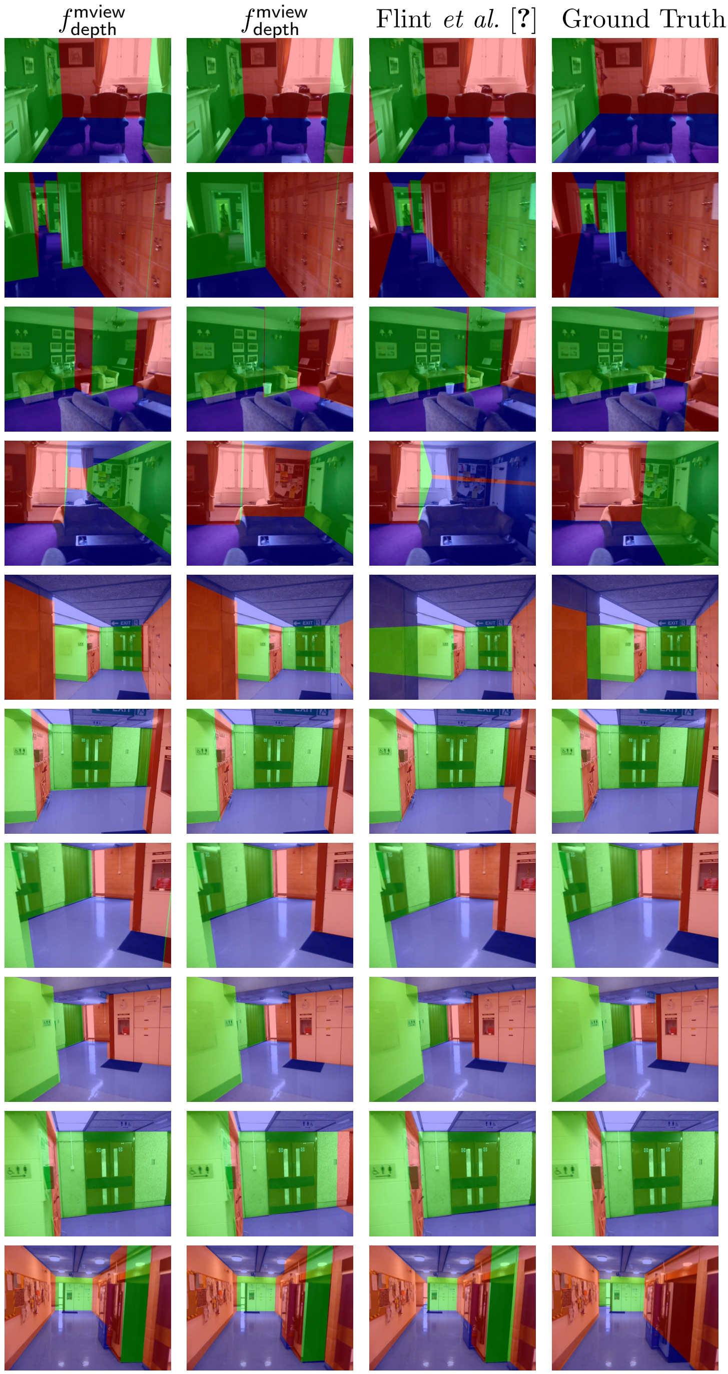
$f^{\mathsf{mview}}_{\mathsf{depth}}$	$f^{\mathsf{mview}}_{\mathsf{depth}}$	Flint *et al.* [?]	Ground Truth



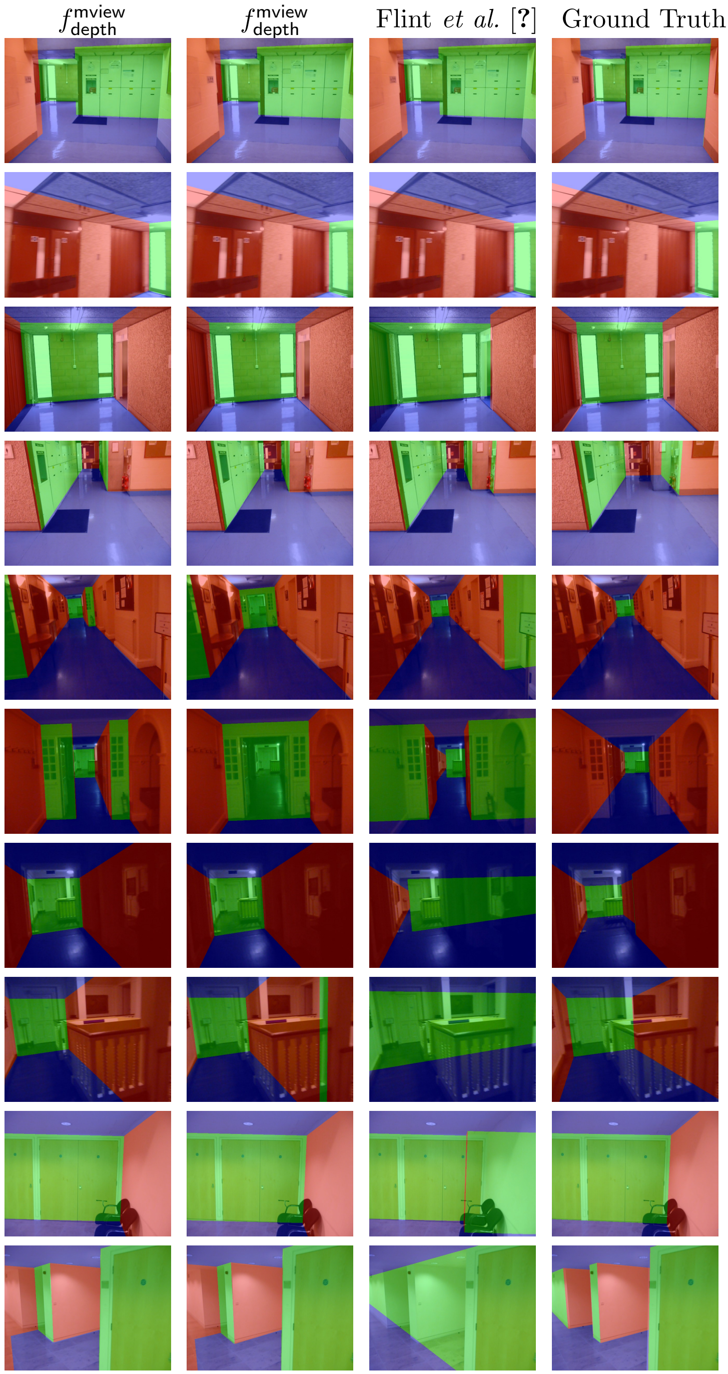
Continued on next page



Continued on next page



Continued on next page



Continued on next page

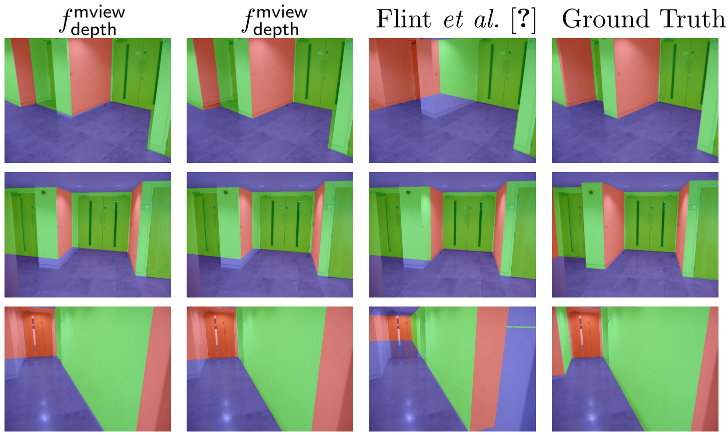
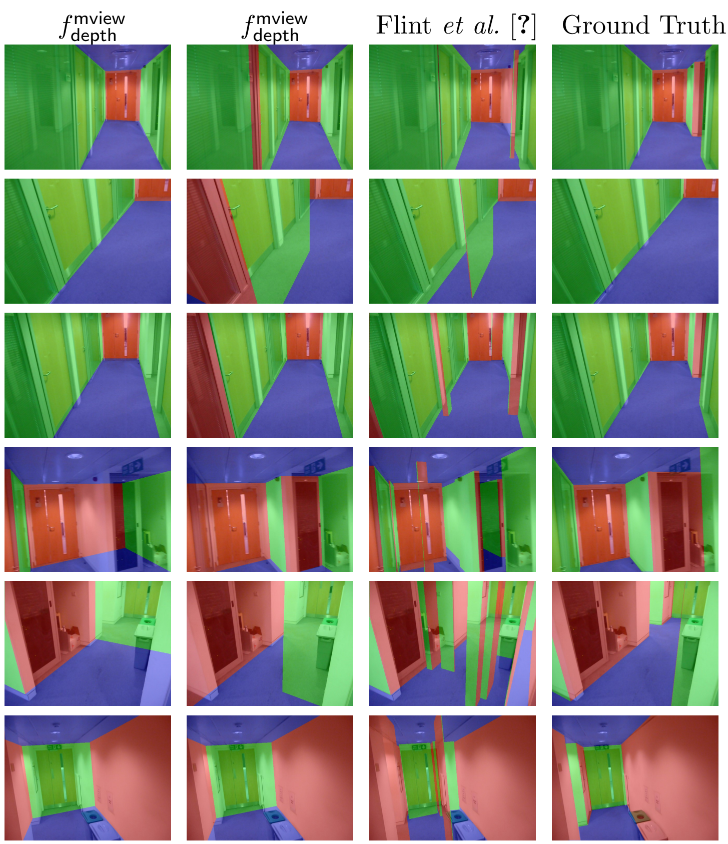


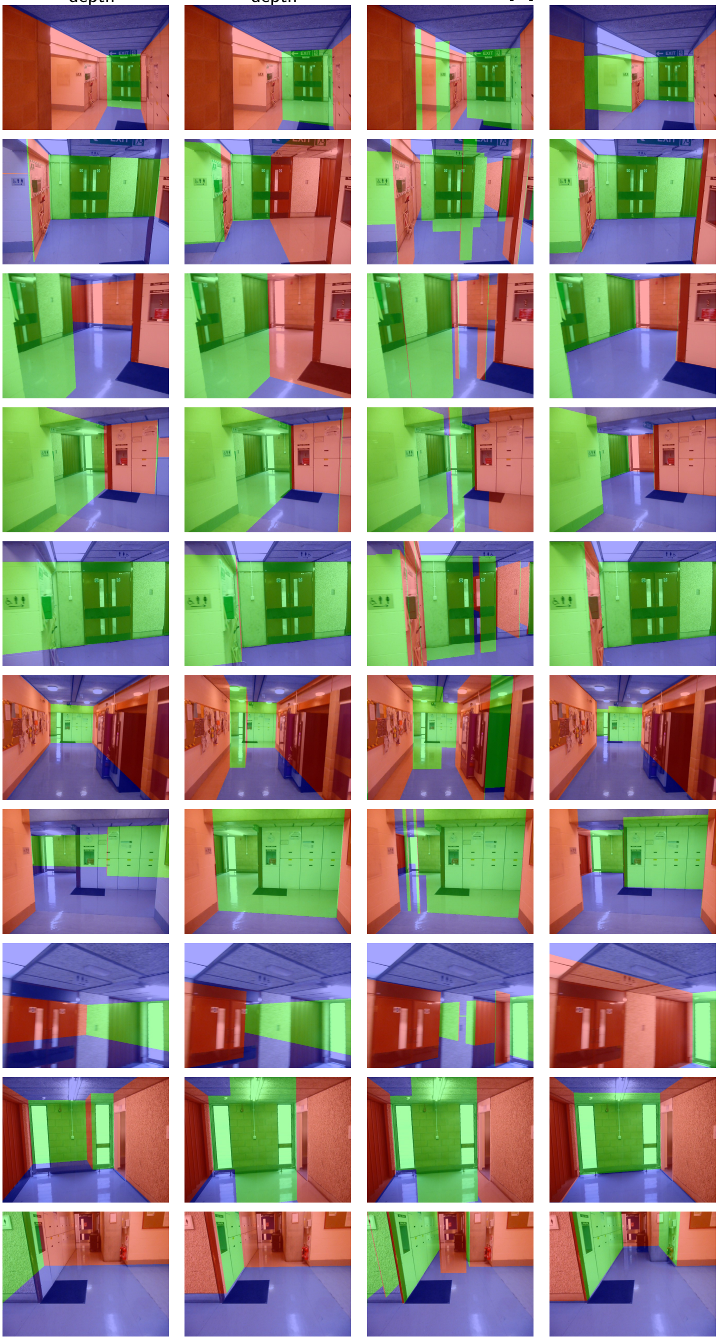
Table 2: Reconstructions predicted by our system using the single view feature space; comparison with Flint *et al.* [?]. Instances in this figure are samples from the hold–out set.



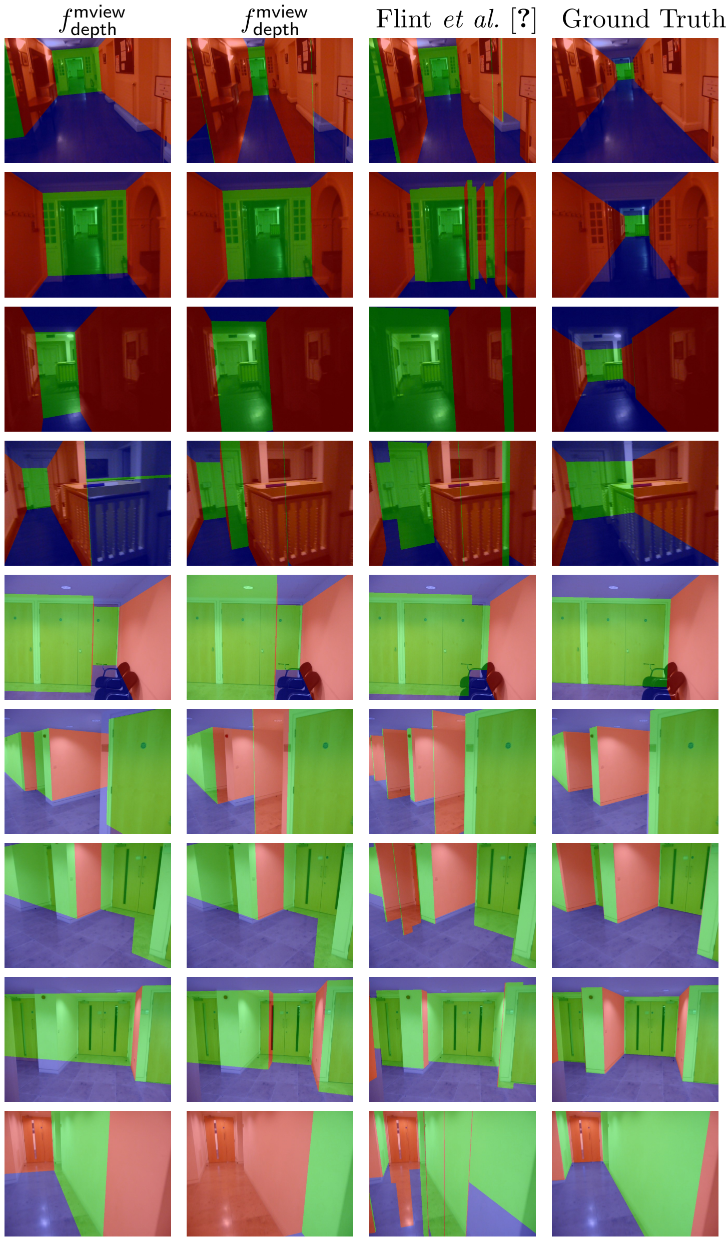
Continued on next page

Continued on next page

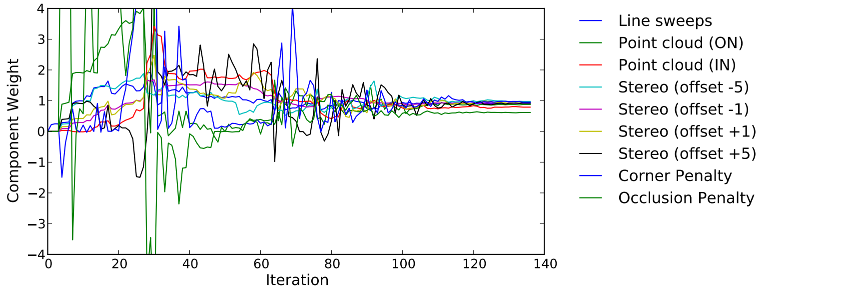
$f^{\text{mview}}_{\text{depth}}$	$f^{\text{mview}}_{\text{depth}}$	Flint *et al.* [?]	Ground Truth



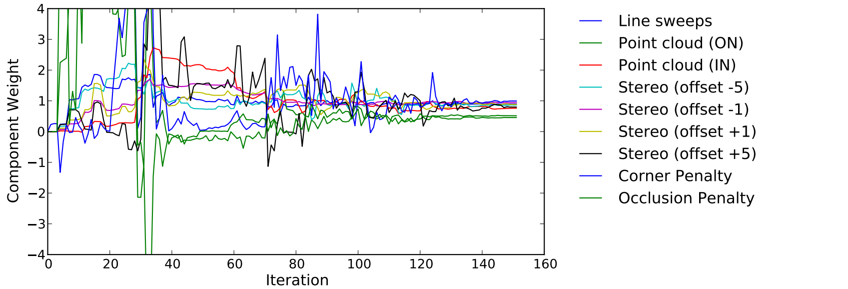
Continued on next page



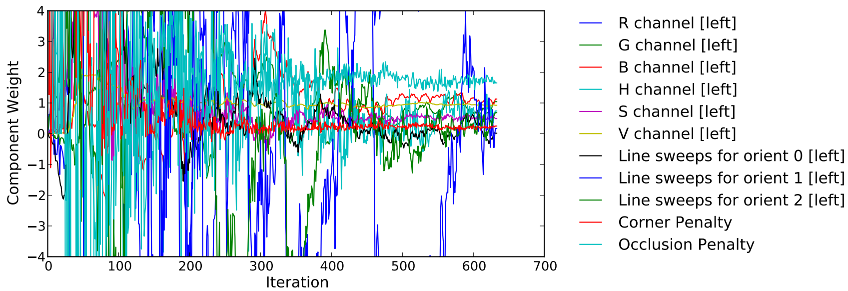
$f^{\text{mview}}_{\text{depth}}$
$f^{\text{mview}}_{\text{depth}}$
Flint *et al.* [?]
Ground Truth



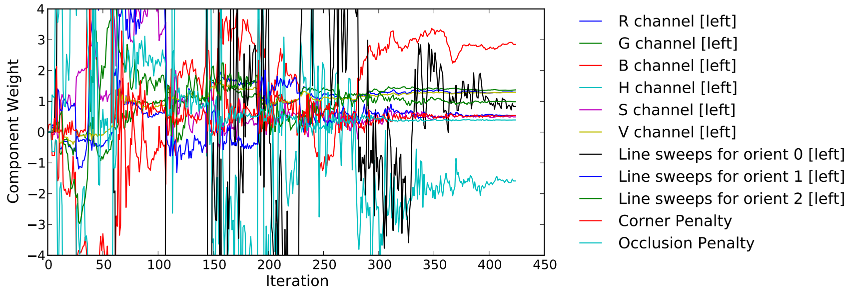
(a) Evolution of $f^{\mathsf{mview}}_{\mathsf{depth}}$ during training



(b) Evolution of $f^{\mathsf{mview}}_{\mathsf{labelling}}$ during training



(c) Evolution of $f^{\mathsf{sview}}_{\mathsf{depth}}$ during training (selected features only)



(d) Evolution of $f^{\mathsf{sview}}_{\mathsf{labelling}}$ during training (selected features only)

Fig. 1: Evolution of weights during training of each predictor.

Sequence	$f^{\mathsf{mview}}_{\mathsf{depth}}$	$f^{\mathsf{mview}}_{\mathsf{labelling}}$	[?]	$f^{\mathsf{sview}}_{\mathsf{depth}}$	$f^{\mathsf{sview}}_{\mathsf{labelling}}$	[?]
`ground`	4.9	5.5	66.6	17.3	28.0	24.5
`foyer1`	6.1	4.5	6.6	25.1	29.2	31.0
`foyer2`	4.3	4.9	5.4	29.1	28.5	30.1
`corridor`	14.6	15.0	52.9	31.7	32.7	33.6
`mcr`	34.0	37.3	67.6	70.1	58.6	45.9
`kitchen`	16.8	19.3	23.6	25.1	25.8	26.2
Average	**13.4**	14.4	37.1	33.1	33.8	31.9

(a) Relative depth error

Sequence	$f^{\mathsf{mview}}_{\mathsf{depth}}$	$f^{\mathsf{mview}}_{\mathsf{labelling}}$	[?]	$f^{\mathsf{sview}}_{\mathsf{depth}}$	$f^{\mathsf{sview}}_{\mathsf{labelling}}$	[?]
`ground`	2.8	2.9	10.4	14.8	7.8	12.4
`foyer1`	3.6	3.1	3.1	20.3	15.1	22.2
`foyer2`	3.3	3.7	4.0	18.6	15.9	18.6
`corridor`	11.4	9.5	19.2	23.7	19.3	24.8
`mcr`	15.4	15.8	16.2	24.5	26.7	20.8
`kitchen`	5.3	5.2	6.1	11.0	7.7	11.9
Average	7.0	**6.7**	9.8	18.8	15.4	18.4

(b) Labelling error

Fig. 2: Performance for each predictor, for each sequence, and for each error metric. The six predictors under comparison are listed in the top row of each table. Note that even when a predictor is trained with respect to a particular loss function, we may still evaluate it on hold–out data using a different metric. However, as exepected the predictors perform best when evaluated by the same error metric that they were each trained for.