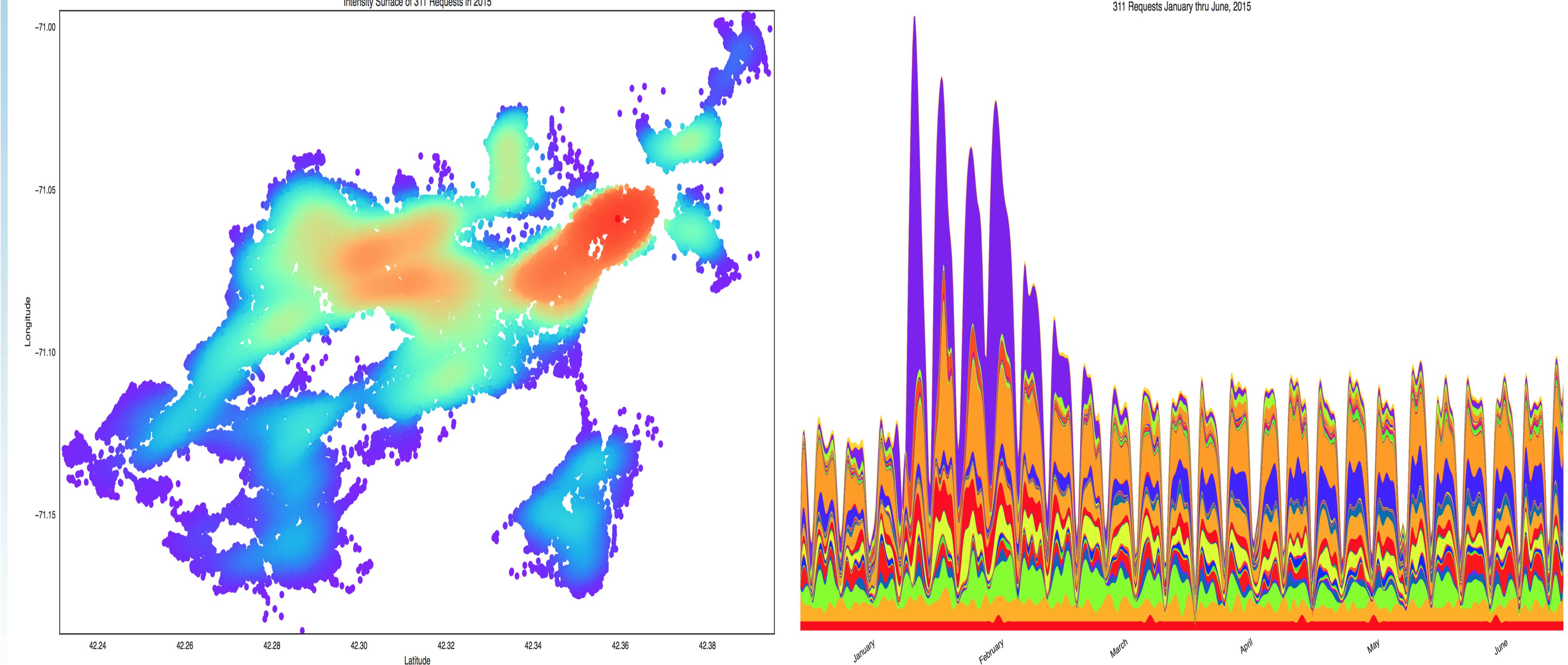
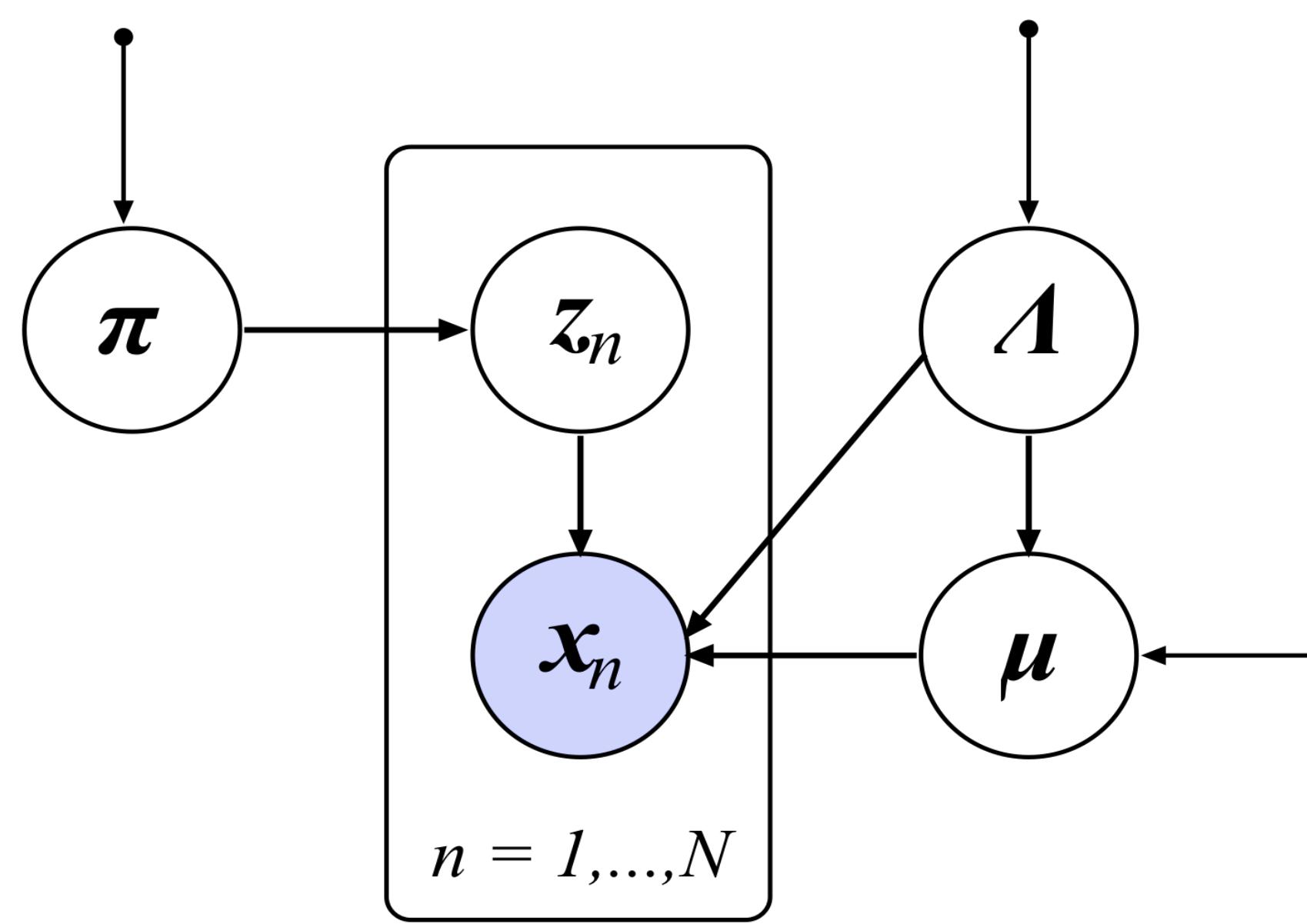


## INTRODUCTION:

Inverse compositional image alignment (ICIA) is fast, but not reliable. We explain ICIA from a different perspective which leads naturally to two new algorithms with a better capture range and comparable speed.



## THE MODEL

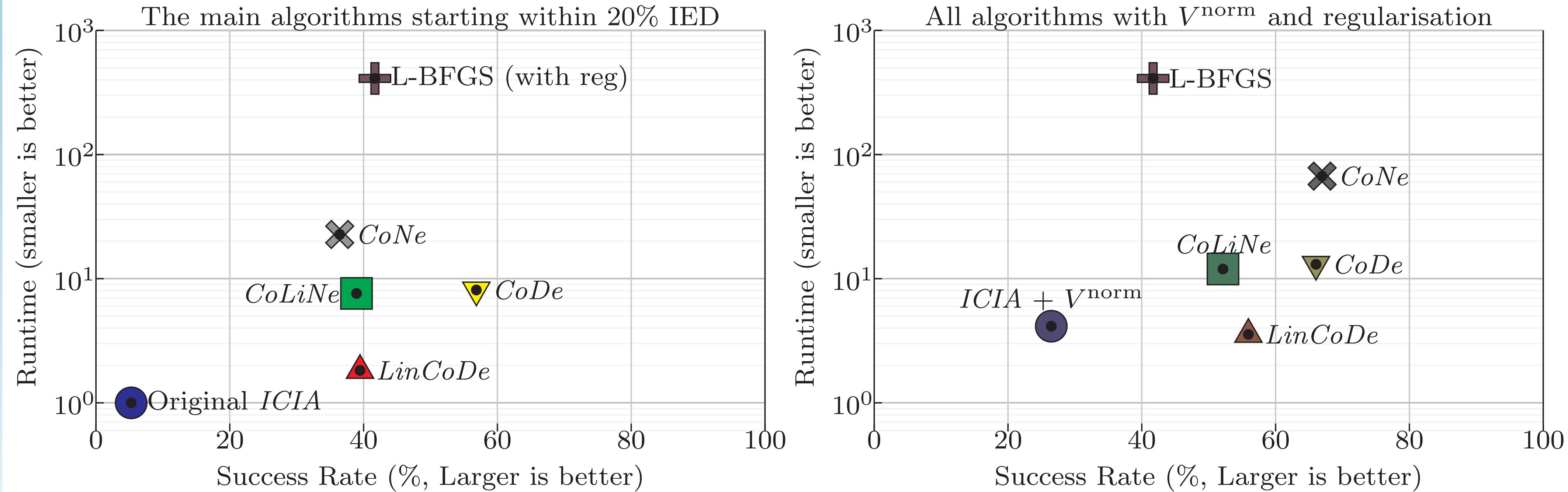


## INFERENCE

```

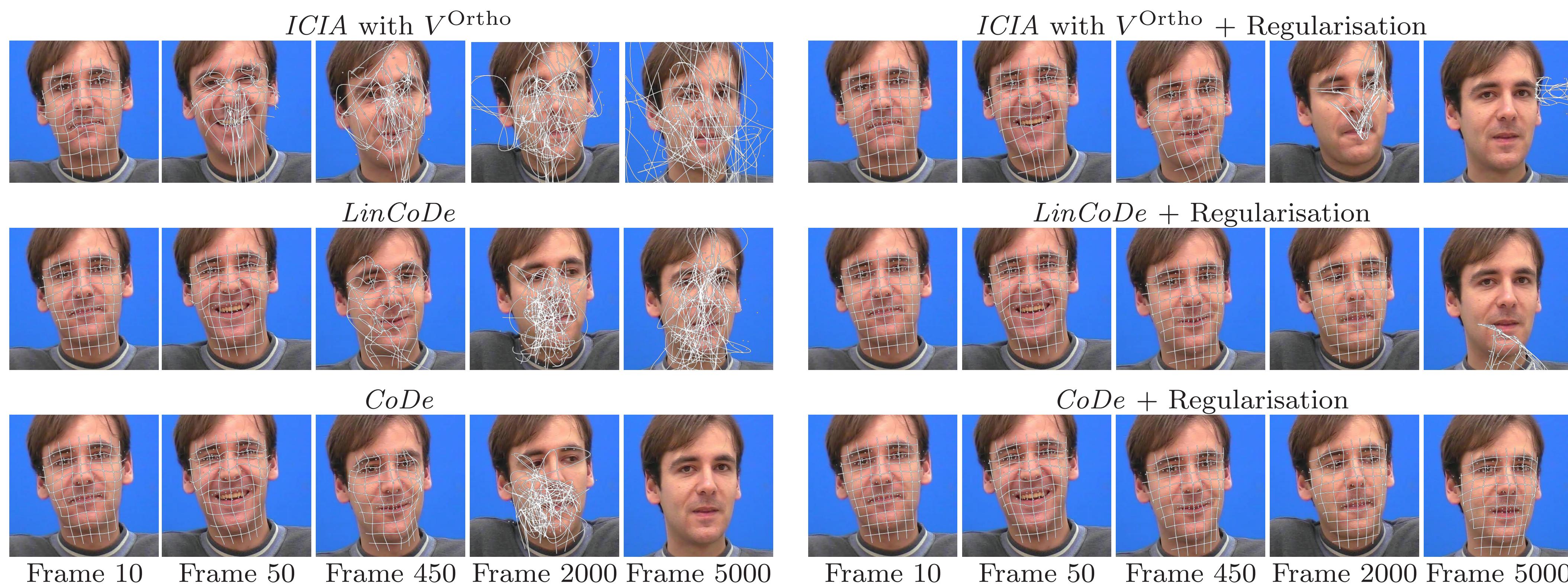
for Blur and regularisation values do
1   Initialize  $\mathbf{q}$ ,  $\mathbf{q}_{best}$  and  $\kappa$ 
repeat
2   Calculate  $\nabla_{\mathbf{p}} \tilde{F}(\mathbf{q}, \mathbf{0})$ ,  $F(\mathbf{q})$ 
3   if  $F(\mathbf{q}) < F(\mathbf{q}_{best})$  then
4      $\mathbf{q}_{best} \leftarrow \mathbf{q}$ 
      Increase  $\kappa$ 
    else
      if  $\kappa$  smaller than threshold then
        return
        decrease  $\kappa$ 
    end if
    Calculate  $\mathbf{p}$  from  $\nabla_{\mathbf{p}} \tilde{F}(\mathbf{q}_{best}, \mathbf{p})$  and  $\kappa$ 
     $\mathbf{q} \leftarrow C^\circ(\mathbf{q}, \mathbf{p})$ 
until converged
  
```

## LATENT CLASSES IN 311 DATA



**Fitting a multiperson AAM.** The best speed–performance tradeoffs come from the two new algorithms *CoDe* and *LinCoDe*. Note that *ICIA* is practically useless on this difficult multi-person dataset with a success rate near zero (left). It can be improved (right) by using the orthonormal incremental warp and regularisation. The *CoDe* algorithm with regularisation (right) is as accurate as the slow, approximation-free, compositional Gauss–Newton *CoNe* method but is seven times more efficient.

## VISUALIZATION



**Our algorithm makes fast and robust tracking possible.** We compare face tracking under natural motion, using *ICIA*, *LinCoDe* and *CoDe*. The original *ICIA* fails immediately with this large model and new face data. Substituting the orthonormal incremental warp for the original *ICIA* warp, the algorithm still loses track very early, whereas *LinCoDe* and *CoDe* can track much further. Finally, adding regularisation to all algorithms, *ICIA* still loses track completely after approximately 500 frames and does not recover the local deformations accurately. In contrast *CoDe* now tracks the full 5000 frame sequence without reinitialization, and *LinCoDe* tracks for 2500 frames.

The same training dataset was used for both tracking experiments. The training data was acquired with different camera and light settings from different subjects.