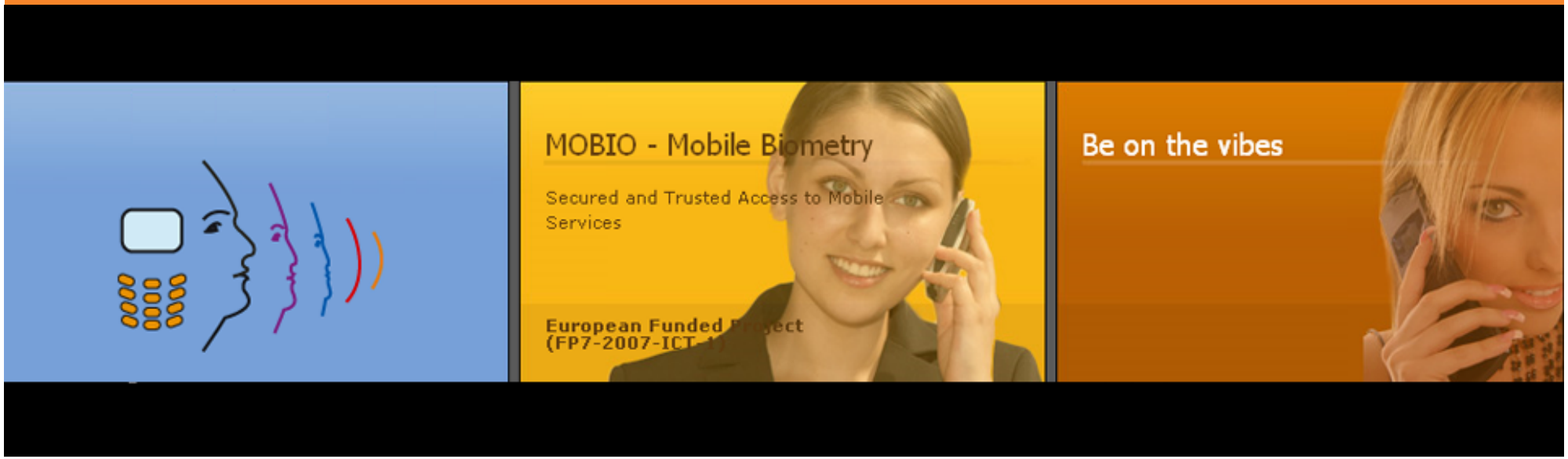
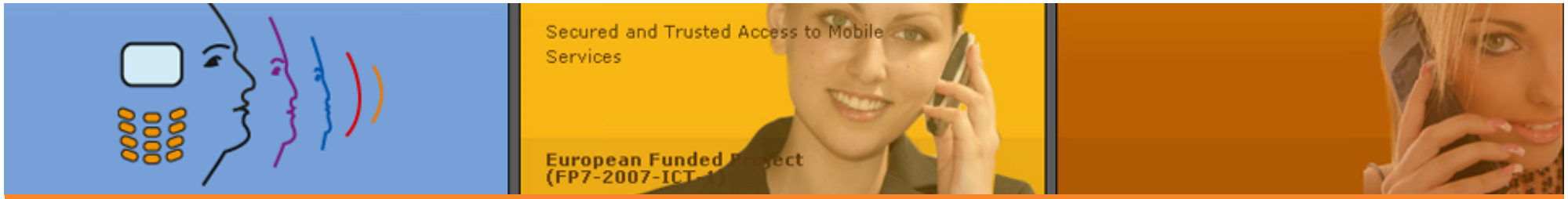
D3.3 Advanced Unimodal System

UMAN

MOBIO Review Meeting, Sep.16-17, 2009

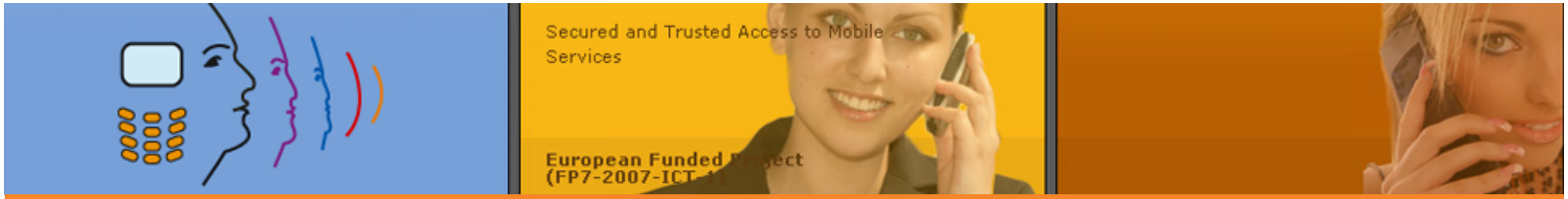
EyePmedia – 1020 Renens





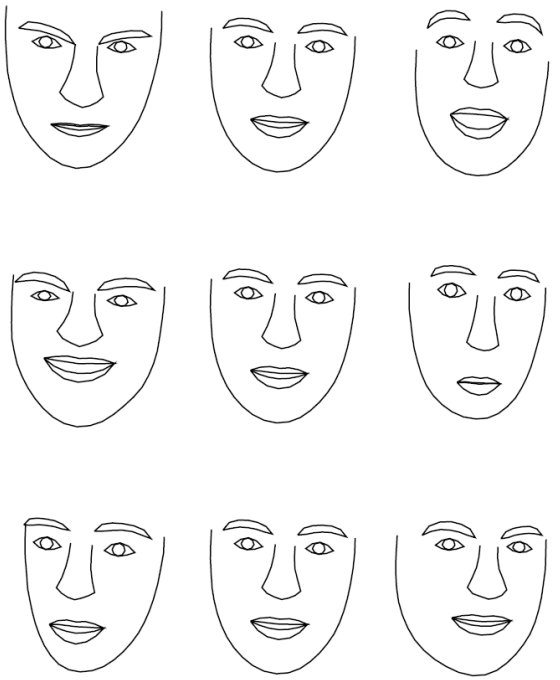
D3.3 UMAN Advanced Unimodal System

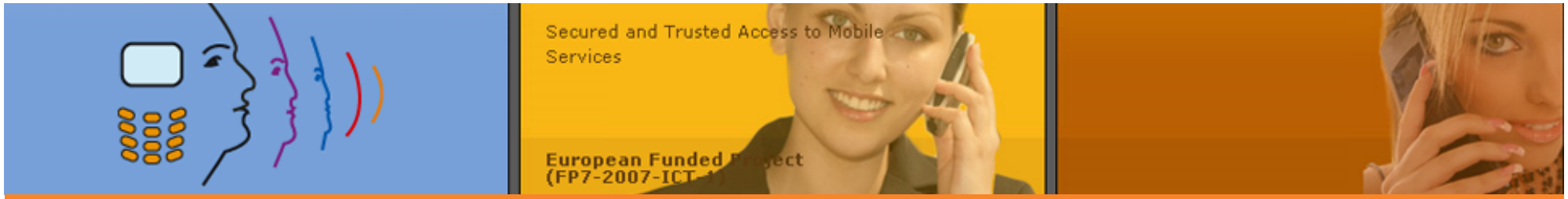
- Recap
  - Shape
  - Appearance
  - Matching
- Guided candidate selection
- Hierarchical implementation
- Results



Recap: Shape

- Photograph lots of people
- Label anatomical features on the face
  - Eyes
  - Nose
  - Mouth
- Compute mean shape and modes of deformation



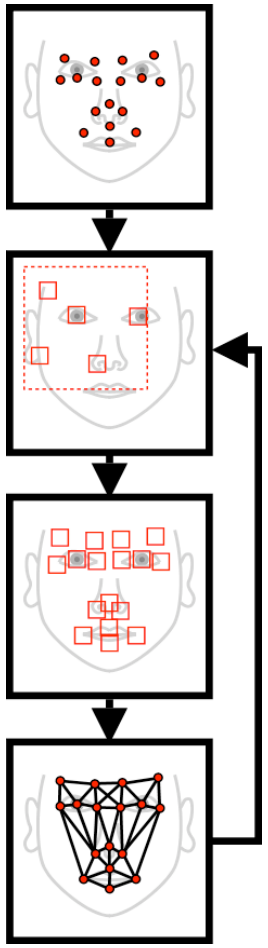


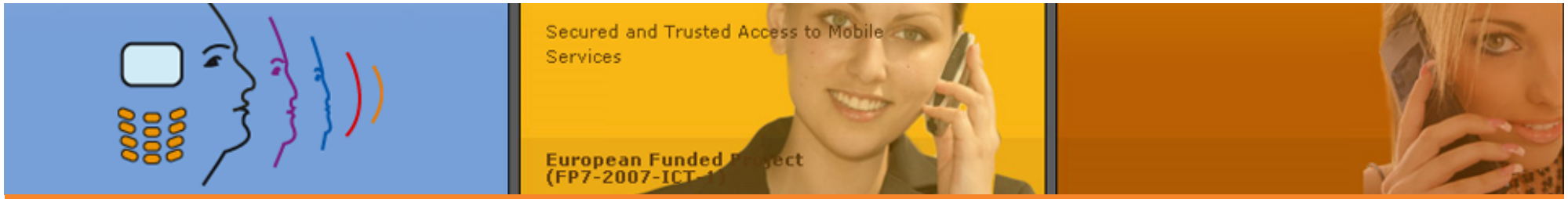
Recap: Appearance

- Photograph lots of people
- Resample image region around each labelled feature
- Compute mean of samples to give template
  - May pre-process image first
  - E.g. compute gradients

Recap: Naïve matching

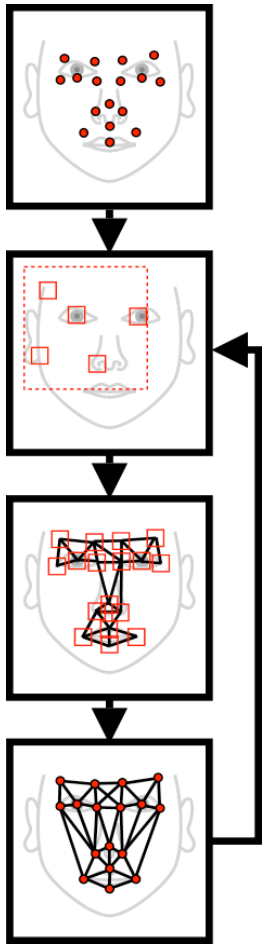
- Guess where features are
  - Mean shape wrt detected face
- Search a nearby region for patches that match template
- Pick best match for each feature
- Project selected candidates onto nearest ‘valid’ face
- Repeat to convergence





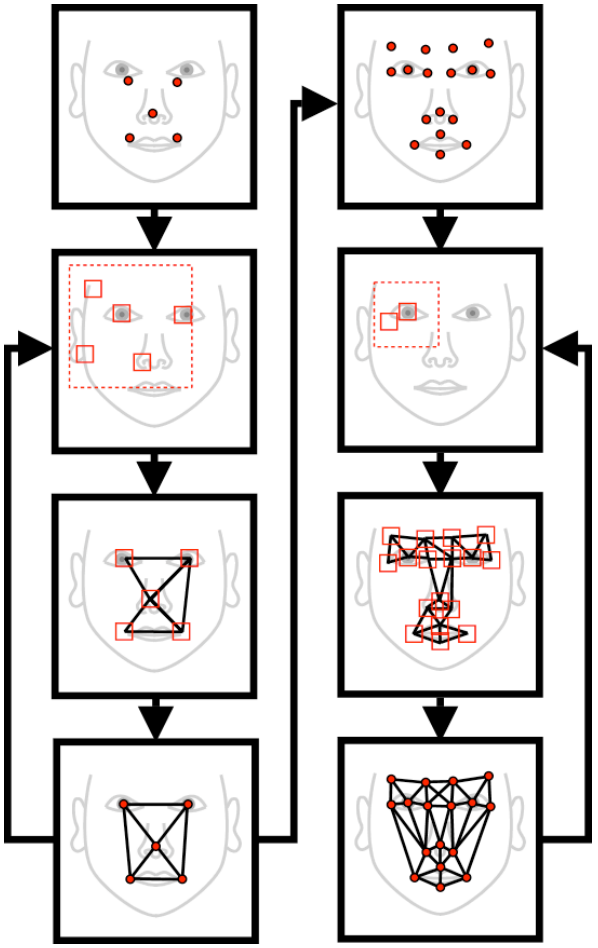
Guided candidate selection

- Best match for each feature selected independently
  - Ignores spatial constraints
  - E.g. both eyes could be assigned to same image patch
- Can be remedied using a Markov Random Field prior
  - Approximates joint distribution
  - Efficient graph solvers (e.g. dynamic programming)



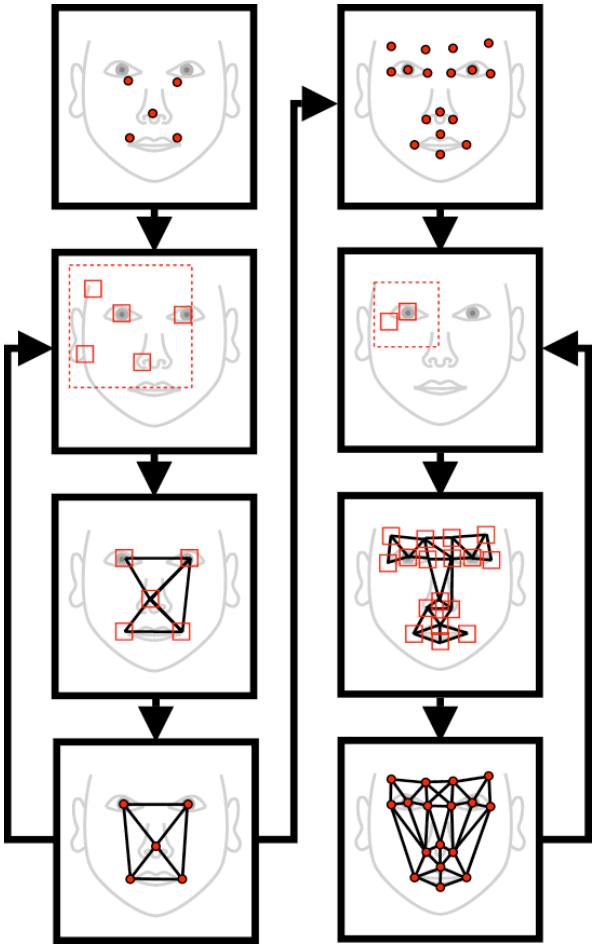
Hierarchical implementation

- Initial uncertainty in scale and orientation is high
  - Need large search regions
  - False matches to undistinctive features are ‘distracting’
- Start with most salient features only
  - Eyes, nose, mouth



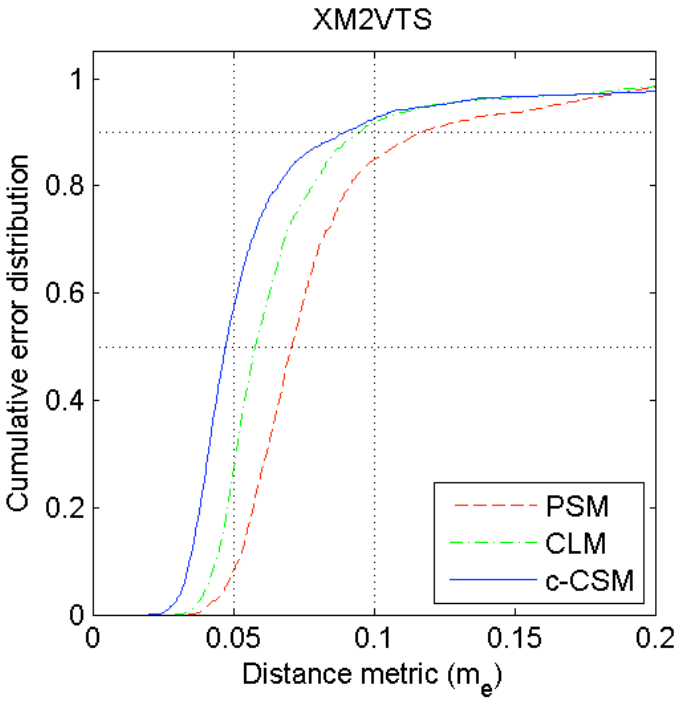
Hierarchical implementation

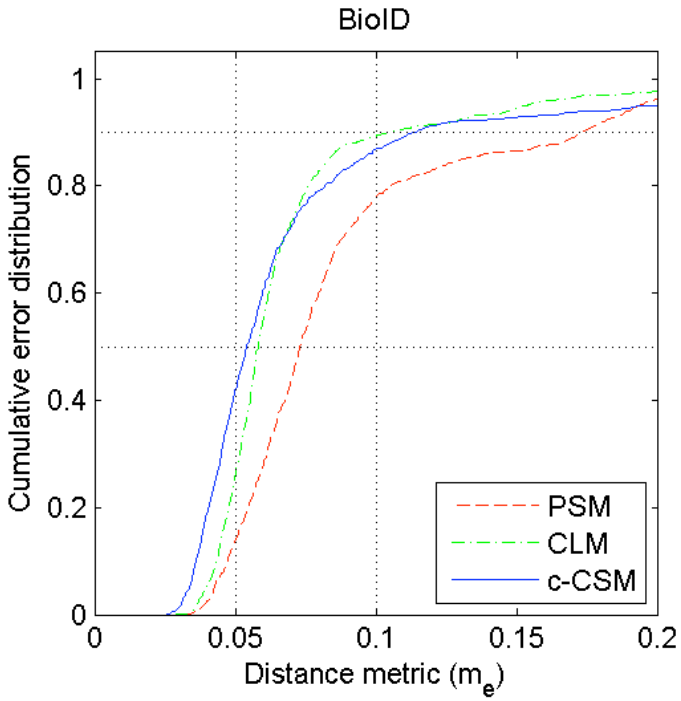
- Uncertainty in scale and orientation gets smaller
- Can make better prediction of other feature positions
- Smaller search regions
  - Fewer false matches
  - More efficient

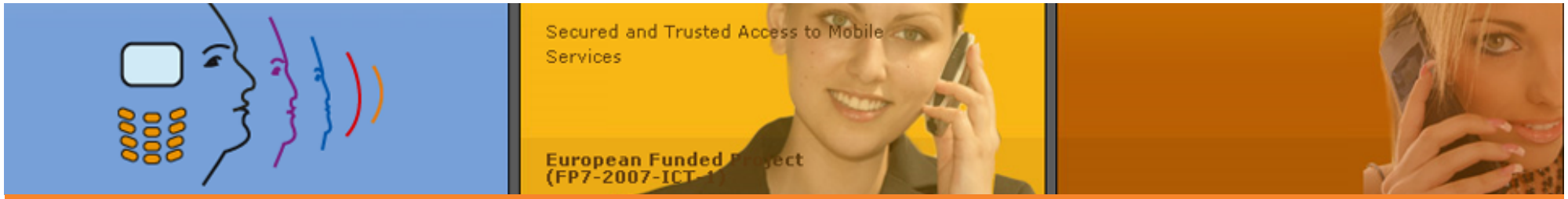


Results

- Tested on standard benchmarks
  - XM2VTS, BioID

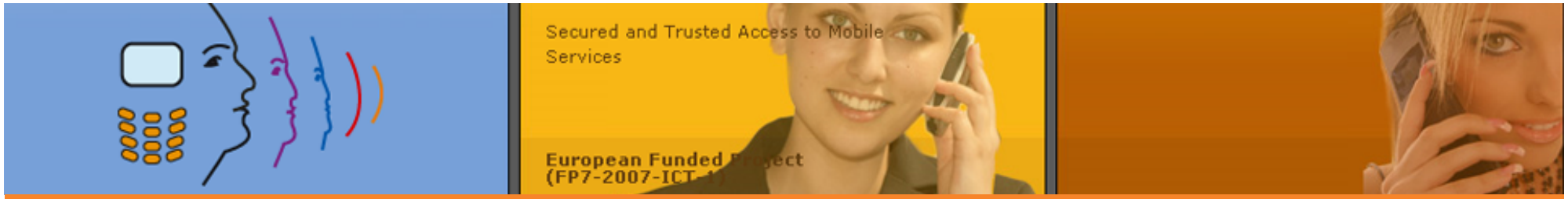






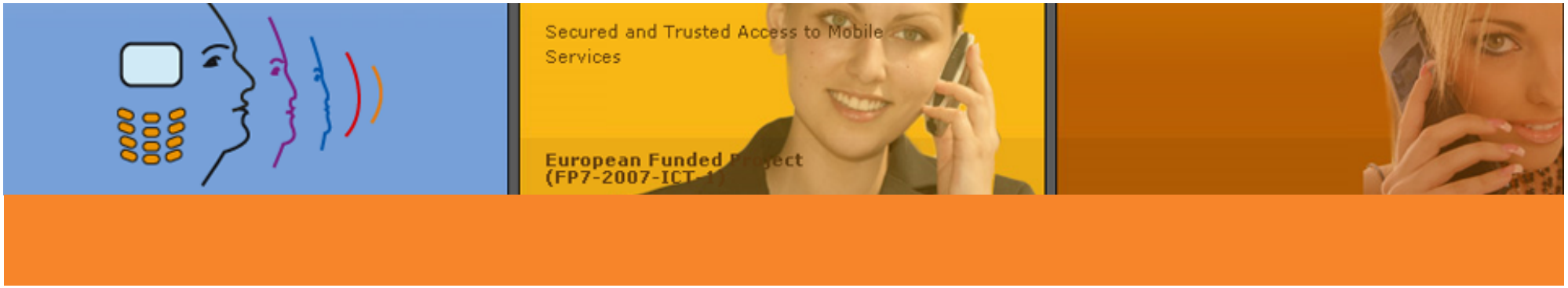
Discussion

- Results compare favourably with previous methods
  - Clearly outperforms baseline on XM2VTS
  - Outperforms baseline on BioID in ~70% of cases
- Recently presented at BMVC (Sep 2009)



Ongoing research

- Improved image representations
  - More accurate detections (i.e. close to true solution)
  - Fewer false matches
- Optimized Markov Random Field structure
  - Maximize accuracy for a given complexity
  - Minimize complexity for a given accuracy
- Exploiting sequence dynamics for efficient tracking of facial features



Thank you for your attention