Computational Photography and Capture:

Video Sprite Animation

Gabriel Brostow & Tim Weyrich

TA: Frederic Besse

Video Sprites

- Web-page
- Local link to Video Sprites summary

Controlled Animation of Video Sprites

Schoedl & Essa, SCA 2002

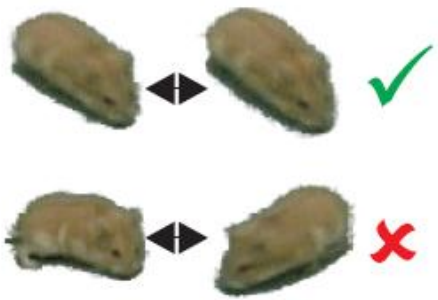
- Optimize animation w.r.t. user-defined costs
- Account for some perspective projection

*Arno Schoedl is now at Think-Cell

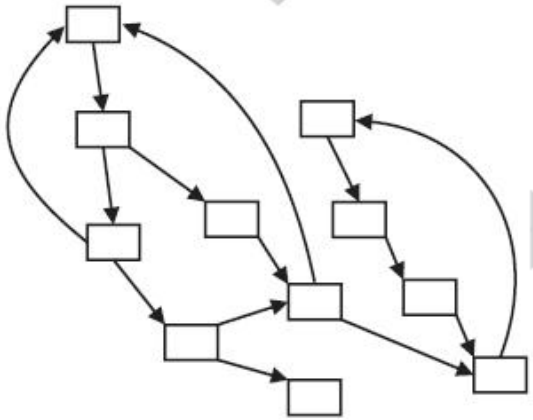
Data capture

Extract sprites
using chromakeying

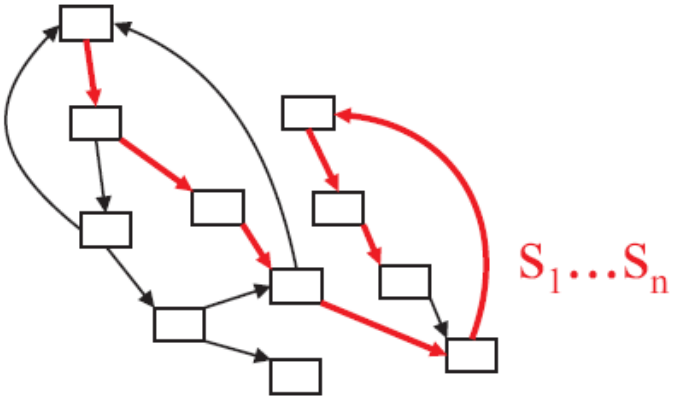
Constraints,
e. g. motion trajectory



List of training pairs



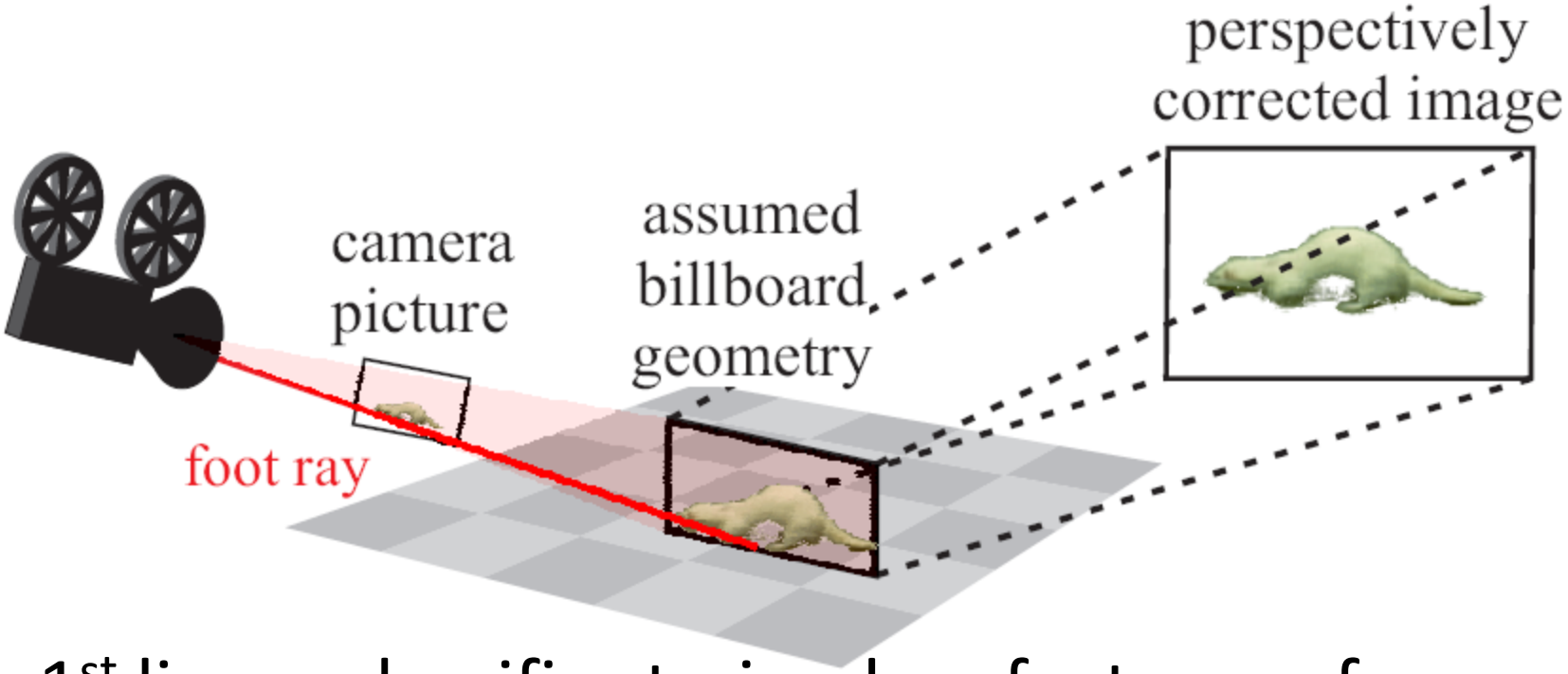
Find transitions by comparing
all pairs of frames



Find sequence of frames $s_1 \ldots s_n$ that
shows desired animation

Render and composite

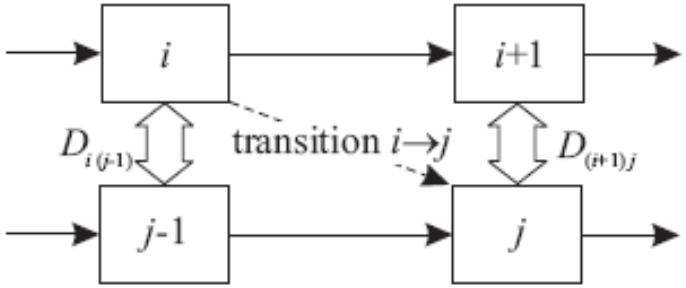
Treat object as billboard in 3D



- 1st linear classifier trained on features of:
  - Sprite velocity, average color, area, eccentricity
- 2nd (as cascade): alpha and color per pixel

Cost Function

- Start with same transition cost on smoothness



- Total cost of frame sequence S (what we're optimizing):

$$C(S) = C_s(S) + C_c(S)$$

$$C_s(S) = \sum_{i=1}^{n-1} C_{s_i \rightarrow s_{i+1}}$$

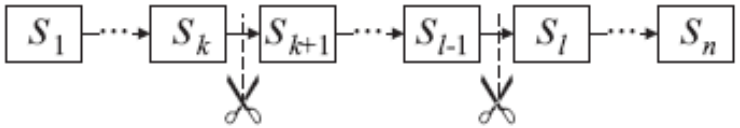
Control Cost Function

Iterated Subsequence Replacement

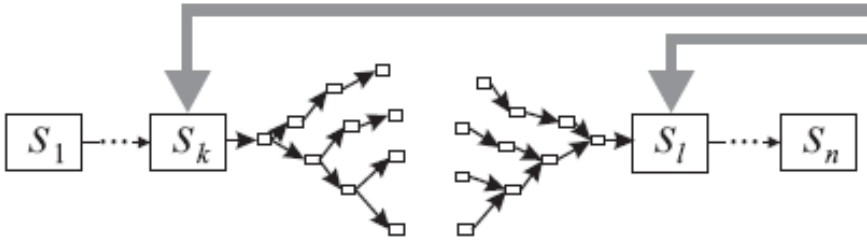
- Q-Learning restricted to overlapping loops and short look-ahead
- Beam-search only adds 1 frame at a time, so no scope for multiple sprites

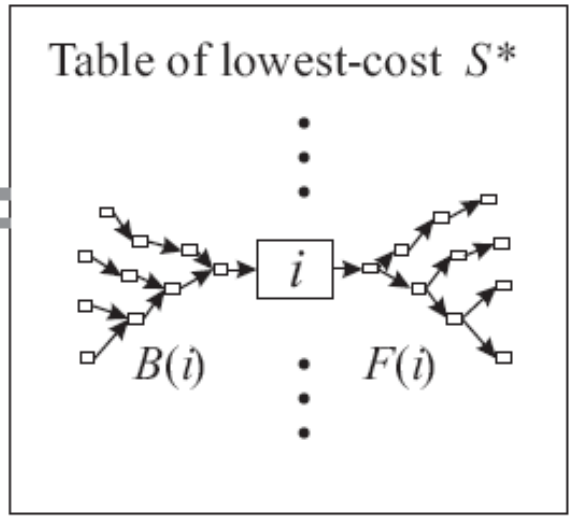
- Must iterate to optimize frame sequences of multiple sprites jointly + for very long look-ahead!
  - Precompute Forward and Backward costs based only on smoothness, every time a subsequence is chosen for replacement

1. Choose randomly s_k and s_l and replace subsequence in between.

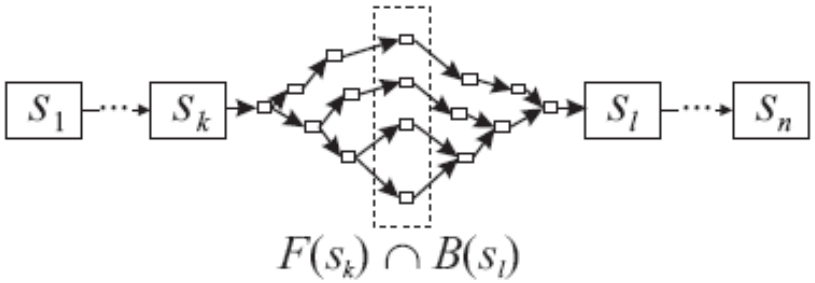


2. Look up precomputed $F(s_k)$ and $B(s_l)$ in path table.

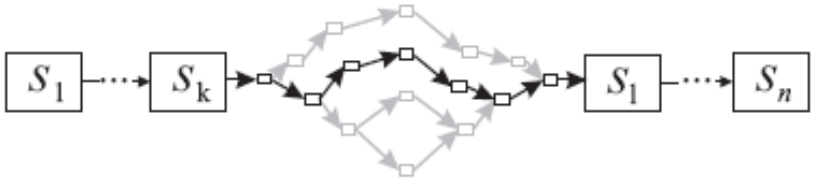




3. Try all $t \in F(s_k) \cap B(s_l)$



4. Keep subsequence for which the total cost improves (if any).



Control Cost C_C

- Sum of costs over all constraints and time steps *i*
- State of sprite at frame *i* is $(p,v,f)_i$
  - *p:* location of the sprite
  - *v:* sprite's velocity
  - *f:* input frame current sprite is copied from
- Example of location constraint: $c_i(p) = \gamma(p - p_{\text{target}})^2$
- Constraint types:
  - Location, Path, Anti-collision, Frame group

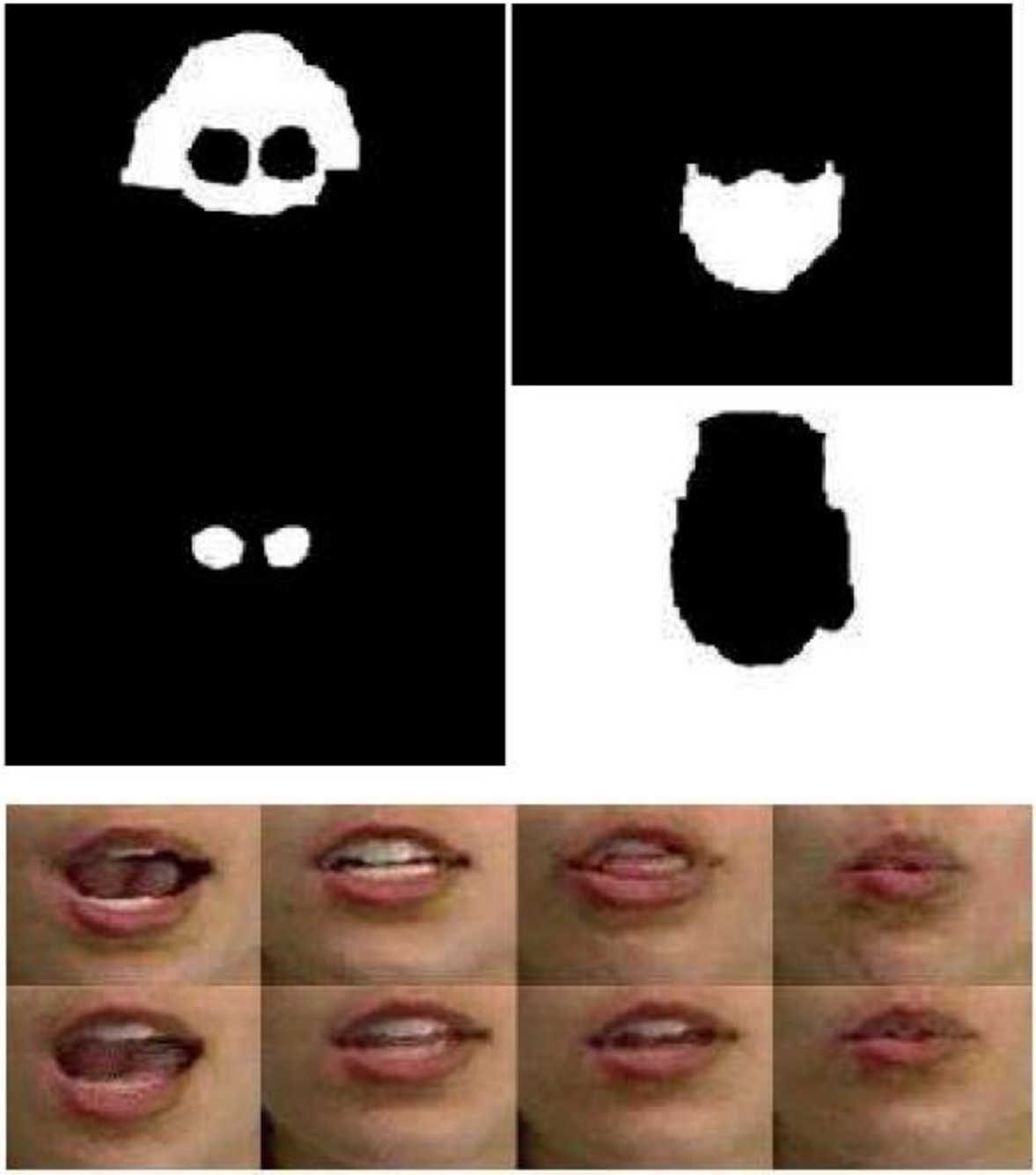
Hamster Path, 2 Locations, Locations & Anti-collision, Location + Group, Formation

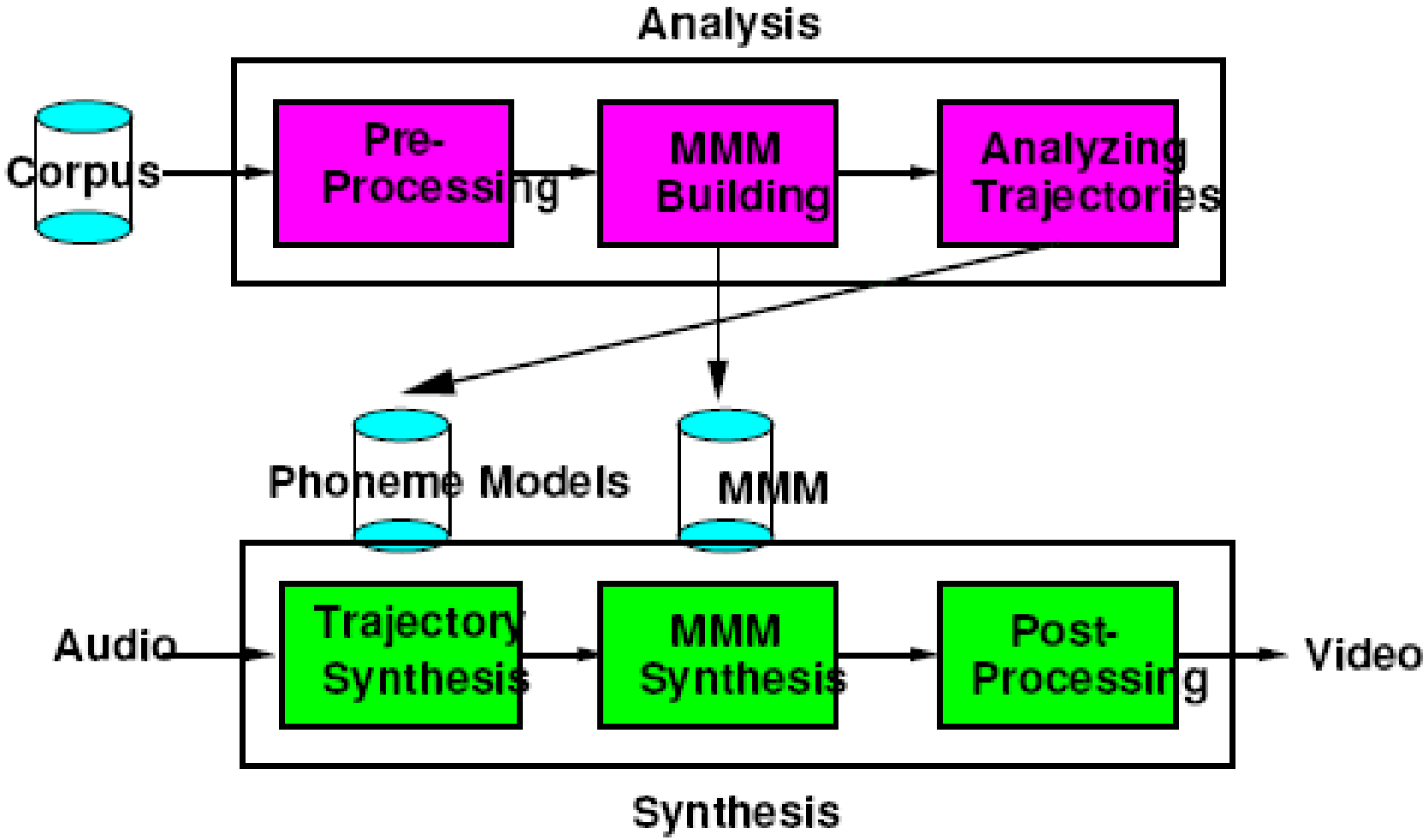
Trainable Videorealistic Speech Animation

- SIGGRAPH 2002 paper by Tony Ezzat, Gadi Geiger, and Tomaso Poggio
  - At Center for Biological and Computational Learning, MIT

- MikeTalk link (from 1998)
- Mary101 (web page of this research)

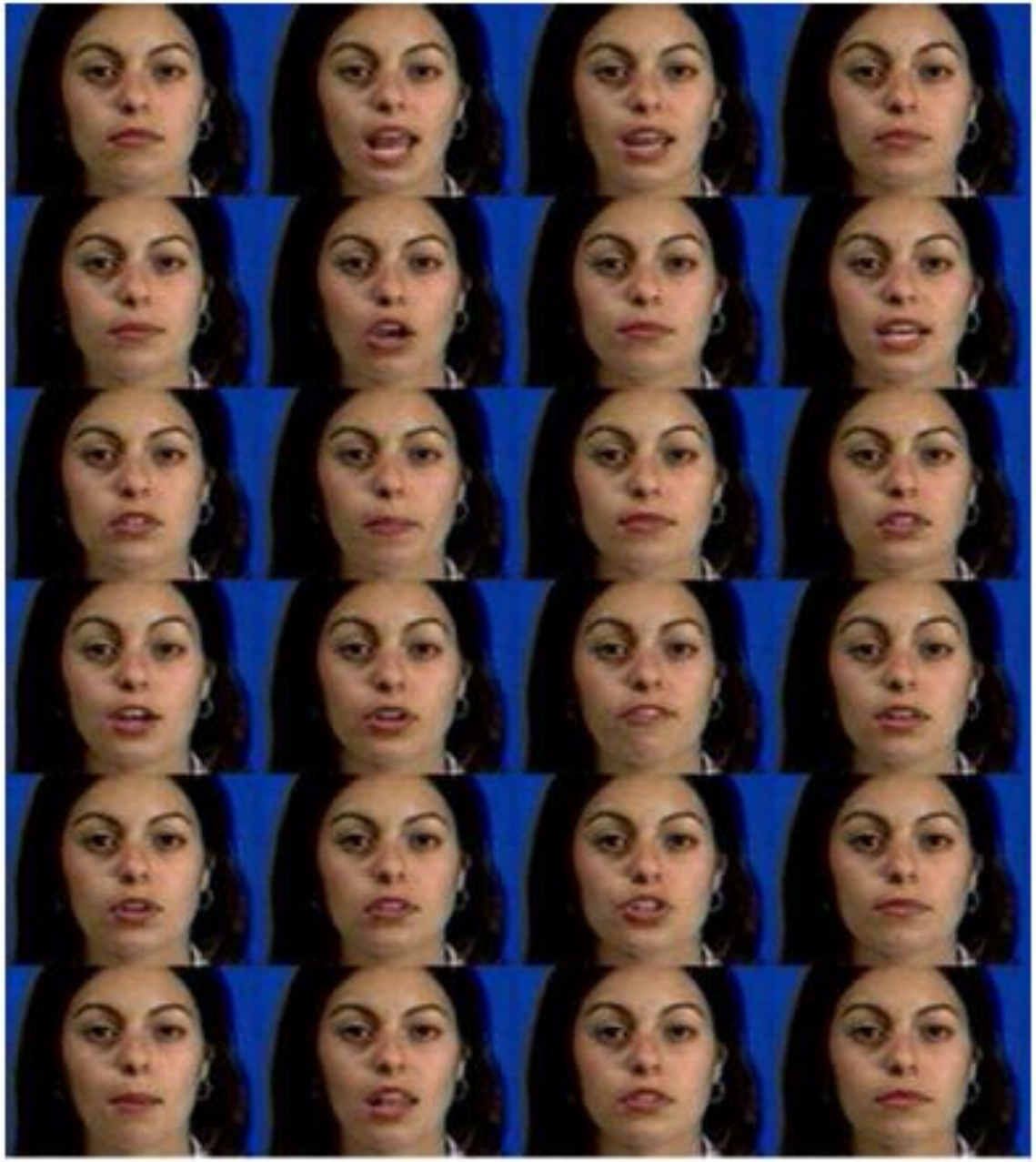
1. Correct motion for the phonemes
2. Smooth transitions
3. Dynamics of plosives (‘b’ and ‘p’)
4. Co-articulation effects




Analysis
Corpus
Pre-Processing
MMM Building
Analyzing Trajectories
Phoneme Models
MMM
Audio
Trajectory Synthesis
MMM Synthesis
Post-Processing
Video
Synthesis

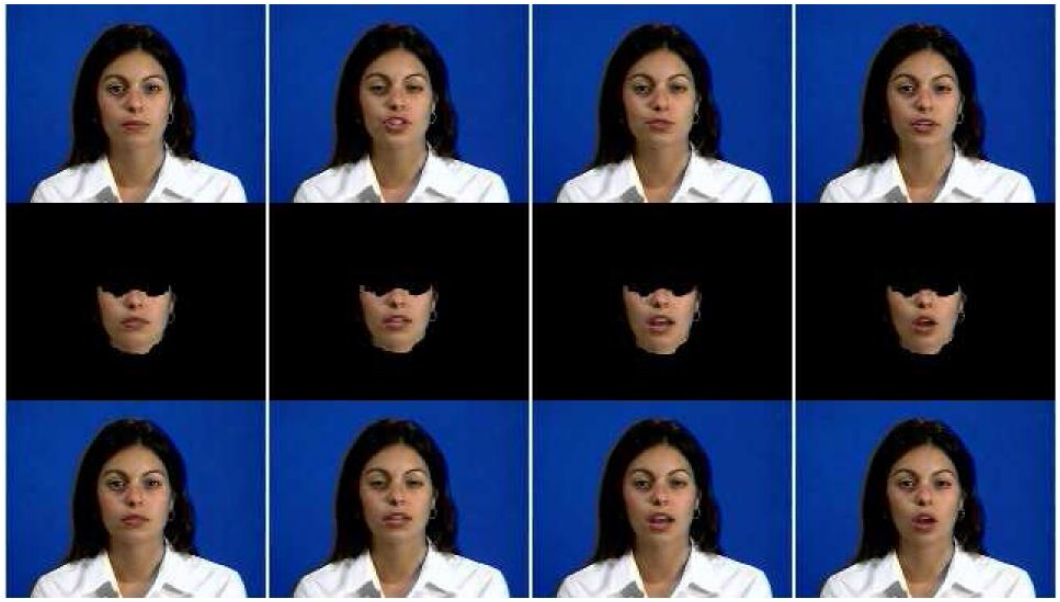
Process

- Stabilize (all 15 min.)
- Phonemes
- MMM: Multidimensional Morphable Model
  - EM-PCA [Roweis98], keep 15 dimensions
  - K-means (N=46)
  - Flow, via Dijkstra on “corpus graph” made with kNN
- Synthesis
  - Trajectory, Render, Composite

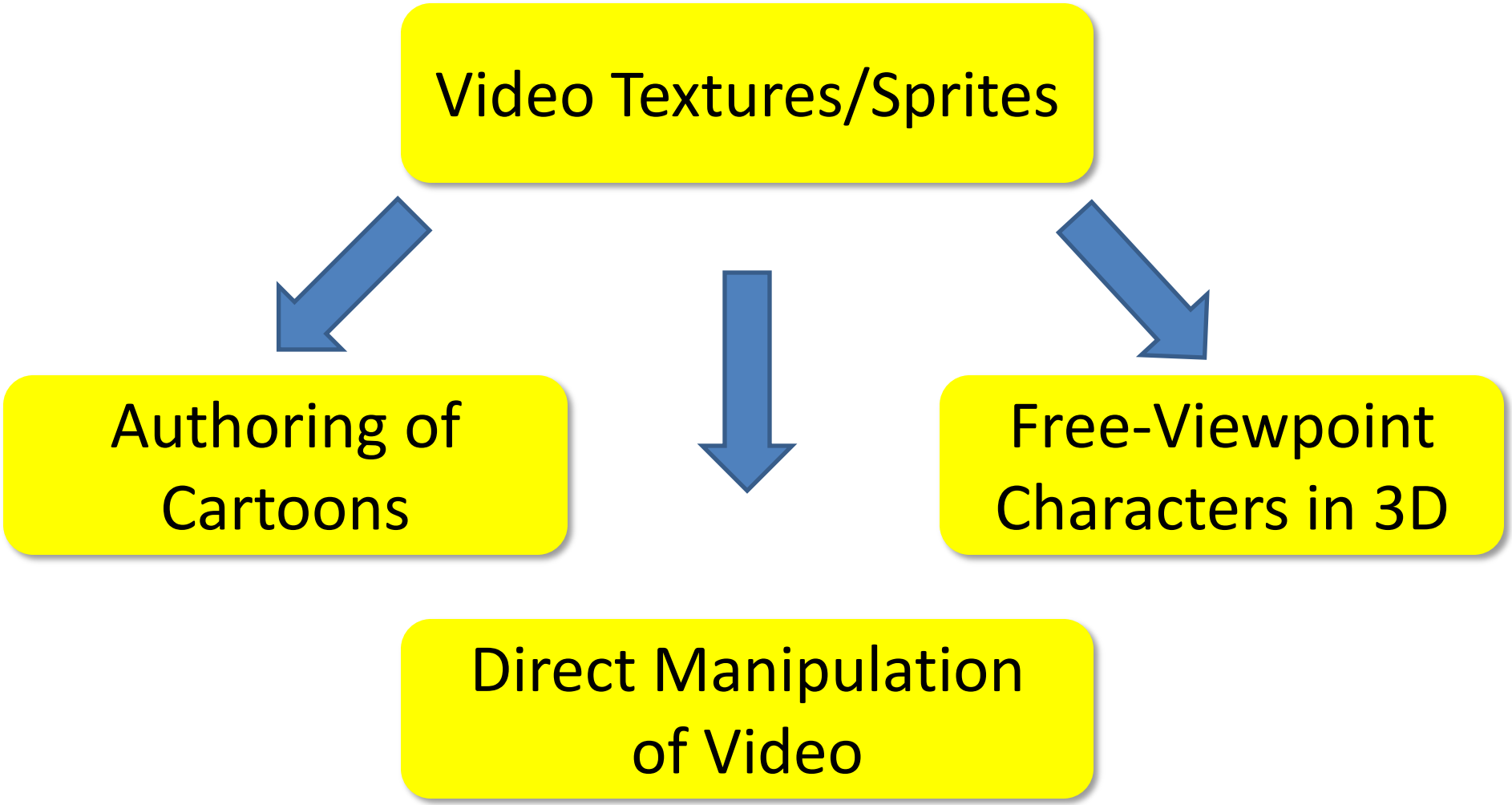


Process

- Stabilize
- Phonemes
- Prototype:
  - EM-PCA
  - K-means
  - Flow (via Dijkstra)
- Synthesis

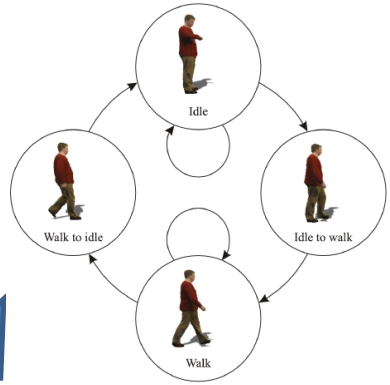


Jump-Off Point to Further Research



Jump-Off Point to Further Research

Video Textures/Sprites



Authoring of Cartoons

Free-Viewpoint Characters in 3D

Direct Manipulation of Video

Video Puppetry, Barnes et al. 2008 v

Cartoon Textures by de Juan + Bodenheimer, 2004

Accessible Animation and Customizable Graphics via Simplicial Configuration Modeling, Ngo et al. 2000

Starck + Hilton:
- Video Based Character Animation v
- de Aguiar et al. 2008
-Vlasic et al. 2008
-Ballan et al. 2008

Direct Manipulation of Video

- DimP: Video Browsing by Direct Manipulation
  - Dragicevic et al., CHI 2008 (v)
- DRAGON: A Direct Manipulation Interface for Frame-Accurate In-Scene Video Navigation
  - Karrer et al. CHI 2008 (v)
- Interactive Video Object Annotation
  - Goldman et al. 2007 (v1-short, v2-long)

- How to map (2D) gestures to object motions?