Extreme Data Streaming Using Image Database

Chun Jia
Voicu Popescu
Problem

- Current limitation in bandwidth for large data communication
- We want to achieve real time interactive communication under low bandwidth network
Key Insights

- Limited body motion
  - Pre-record the motion sequence
  - Send the processed movie in the beginning
- Live video captured from webcam is matched with the frames in image database, only the index of the matched frame is sent
Possible application

Distance Learning Classroom System
Another example

Contribution

- Extreme bit-rate reduction (35kbps -> 480 bps for 30 fps)
- Independent resolution
- Less overhead – no need to build a 3D model
- Low computational complexity: based purely on image comparison
Data Acquisition

- A movie sequence consists of various upper body gestures.
- Natural pausing between sequence.
- Guiding audio instruction.
Data Registration

- Background subtraction
- Frame Segmentation
- Shape Analysis
  - Noise removal
  - Silhouette
Central line (red) and shoulder line (blue) defines the stationary region without shape variation
Model Segmentation

- Group division based on pausing
- Skipping redundant frames
- Merging group based on similar shape
- Binary tree built based on shape variation w.r.t. the stationary region
Tree Visualization

Y
has left shape test

N

Y
has right shape test

N

Y
has right shape test

N
Tree Leaves

- Leaf node 5: left side only shape
Image-based Matching

- Linear searching on all images
- Linear searching on selected tree leaves
- Linear searching on groups
- Binary searching on groups
Linear Searching On All Images

- Use segments in database images with the input image

Optimization
  - Skip frame if bbox not match
  - Use coherence to reduce popping
Linear searching on all images
Linear Searching On Selected Tree Leaves

- Walk down the binary tree and find the matched leaf based on shape
- Searching all images under the matched leaf
  - Can do better if using more efficient matching algorithm in linear searching for all images
  - Further divide the frame into top, left, right region w.r.t. the shoulder line
Subdivision

- Find the discrete curve of the top-, left-, right-most pixel position in each group
- Match the input frame needs only 1D searching
Linear searching on selected tree leaves
Performance

- Current database
  - 1300 frames, captured at 10 fps
  - 500 stored frames, 52 groups, 24 merged groups
Performance Comparison

- Linear searching on all images
  - 68.12 ms per input frame, or 14 fps, use coherence
  - One input frame needs 312 ms, or 0.23 ms computation time for each database image

- Linear searching on selected tree leaves
  - 123 ms per input frame, or 8 fps, no coherence
  - One input frame takes 219 ms computation time
Future Work

- Acquire data using higher resolution
- Capture facial expression
- Build database using multiple people and do matching