Semantically Linking Instructional Content

by Yekaterina Kharitonova

ISTA 352 / University of Arizona / October 5, 2012

Main Goal

To split-up a video into semantically meaningful clips based on the slide usage.

Main Task

Align a video to the slides used during the presentation.

Video: a sequence of video frames.

Slides: a JPG image for each slide.
What is a slide?

1. **Feckless**: generally incompetent and ineffectual
2. **Foible**: a behavioral attribute that is distinctive and peculiar to an individual
3. **Fulmination**: thunderous verbal attack
4. **Gaucherie**: a socially awkward or tactless act
5. **Homiletics**: the act of preaching
6. **Imbroglio**: an intricate and confusing interpersonal or political situation

What is a slide?

1. **Feckless**: generally incompetent and ineffectual
2. **Foible**: a behavioral attribute that is distinctive and peculiar to an individual
3. **Fulmination**: thunderous verbal attack
4. **Gaucherie**: a socially awkward or tactless act
5. **Homiletics**: the act of preaching
6. **Imbroglio**: an intricate and confusing interpersonal or political situation

Need a representation of a slide, which can be “taught” to a computer?
SIFT keypoints

- **Scale-Invariant Feature Transformation**
- Image keypoints - “interesting” points
  - (x,y) position on the image
  - Scale and Orientation
  - 128-dimensional descriptor (texture around the keypoint)

Images courtesy Kobus Barnard and Quanfu Fan
Nearest Neighbor Match

- Match keypoints based on the descriptor

Mapping the keypoints

- Mappings of points on a plane in 3D satisfy a simple relation (linear equation in homogeneous coordinates):

\[
\lambda' \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}
\]

Frame keypoint location (x', y') Homography, H Slide keypoint location (x, y)

Mapping the keypoints

- In homogeneous coordinates, the slide-to-frame mapping is \( X' = H X \)

- The homography matrix, \( H \), is a 3x3 matrix with 8 degrees of freedom

- Frame-to-slide mapping is similar: \( X = H' X' \), where \( H' \) is the inverse slide-to-frame homography

Can all these keypoint matches be correct?
Estimating the homography

- **RANSAC**: **RAN**dom **SA**mple **C**onsensus
- An iterative algorithm that estimates parameters of a model from a set of observed data (which contains outliers).

**Model**: homography transformation

**Model parameters**: \( H \) matrix values

**Observed data**: matched keypoint locations

---

**RANSAC**

Loop
- Select a set \( (S_i) \) of random points
- Fit a model \( (H_i) \) to \( S_i \)
  - If the model is good enough
    - Test all points against the model
    - Keep the best \( H \)
    - Re-estimate \( H \) using all points

Without RANSAC (Nearest neighbour)

Matches limited to \( X = HX \) for some \( H \)
Homography Uses

- Backprojection
- Slide magnification
- Laser pointer-based bullet point magnification

Backprojection

Differences due to:
- Wild animals were leaner
- Foods contained relatively high levels of polyunsaturated fatty acids
- Domesticated animals now fed grains which have high omega-6 levels, low omega-3 levels
- Thus changed composition of chicken, pork, milk
- Also true for farmed fish, since they are fed grain

Slide Magnification

Images courtesy Andrew Winslow and Kobus Barnard
Overview

• Motivation - quickly finding a relevant snippet of the video
• SIFT Keypoints
• Homography
• RANSAC
• Exploiting Homography

Challenges

• Animations (movies, cascading sequences of bullets)
• Blurry slides (out of focus, screen is brushed)
• Speaker occlusion
• Color shifts

Contributors

Arnon Amir (IBM Almaden), Kobus Barnard, Troy Bowman, Joe Chitwood (KUAT), Alexander Danehy, Benjamin Dicken, Alon Efrat, Quanfu Fan, Sandiway Fong, Steven Gregory, YK, Derek Leverenz, Vivek Kumar, Ming Lin, Daniel Mathis, Adam McFarlin, Miguel Solano, Ranjini Swaminathan, Javad Taheri, Mohan Tanniru, Michael Thompson, Juhani Torkkola, Qiyam Tung, Gabriel Wilson, Andrew Winslow, Steve Zhou
Questions?