

SIGGRAPH Content Formatting Instructions



Figure 1: Spring Training 2009, Peoria, AZ.

Abstract

Keywords: Video summarization, Video navigation, Lecture videos, Blackboard-style lectures

1 Related Work

Navigation interface:

- [Kim et al. 2014a] Lecture scape interface
- [Pavel et al. 2014] Video Digest

Directly related to blackboard-style lectures:

- [Monserrat et al. 2013] Clickable visuals and panoramic frame
- [Choudary and Liu 2007] Panoramic frame

Other video summarization

- [Jackson et al. 2013] Panopticon
- [Barnes et al. 2010] Video Tapestry
- [Ekin et al. 2003] soccer video summarization

Comic book style layout of keyframes

- [Boreczky et al. 2000]
- [Uchihashi et al. 1999]
- [Hwang et al. 2006]

Improving readability of video subtitles

- [Hu et al. 2015] optimizing subtitle placement by speaker recognition

- [Kurlander et al. 1996], [Chun et al. 2006] subtitle placement for sinle frame in cinema comics

2 Method

2.1 Time-stamped Transcript

Several on-line video lectures (e.g. Khan Academy) come with transcripts. In cases where transcripts were not provided, we used an on-line audio transcription service to acquire a verbatim text transcript. Then, we used a tool from [Rubin et al. 2013] to compute an alignment between the video's audio file and the transcript. The final output is a time-stamped transcript, where each word is annotated with a start and end time.

2.2 Stroke Extraction

A *stroke* is defined as the set of foreground pixels that is drawn during one continuous drawing activity. The method used to extract strokes from video frames are similar to that used by [Monserrat et al. 2013] to extract visual objects in their NoteVideo interface. Figure 2 shows examples of extracted strokes from different videos. A typical stroke comprises several characters to several words, or it can also be a part of other drawings such as a graph (Figure 2c).

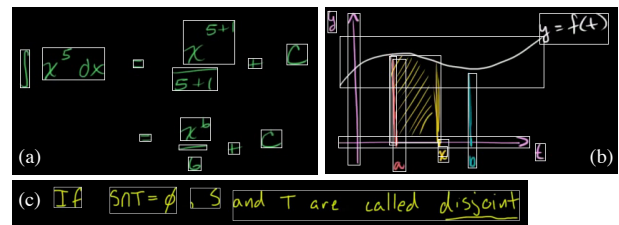


Figure 2: Examples of strokes extracted from different videos.

2.3 Hierarchical Grouping of Strokes

We group strokes into hierarchical units: lines and sentence-strokes. A line consists of a set of strokes that *belong together* semantically. For example, a line could be a single row of equations, or a graph including its labels. Figure 3 shows examples of lines. The problem of grouping strokes into lines is analogous to the problem of line breaking, also known as word wrapping [Knuth and Plass 1981]. An important difference is that in the traditional word wrapping

problem, only a contiguous set of words can be put in the same line. In our case, strokes in a single line can be interspersed by strokes in a different line. This happens, for example, when the instructor goes back and forth between two lines of equations, or between a graph and an equation (Figures ??).

Scoring function description

Pseudo code figure

Input : list of strokes S

Output: list of optimal lines, $L_{|S|}$

$L_{-1} = \{\}$ // L_i = optimal set of lines upto i -th stroke

```

for each stroke  $s_i \in S$  do
  minscore =  $+\infty$ 
  for  $j \leftarrow -1$  to  $i - 1$  do
    for  $n \leftarrow 0$  to  $|L_j| + 1$  do
      // score to merge  $s_i$  to  $n$ -th line of  $L_j$ 

      // If  $n = |L_j| + 1$ ,  $s_i$  on a new line by itself.
      score  $\leftarrow$  line_score( $L_j, n$ )
      if score  $<$  minscore then
        optj =  $j$ 
        optn =  $n$ 
      end
    end
  end
   $L_i$  = merge  $s_i$  to optn-th line of  $L_{optj}$ 
end

```

Sentence is a meaningful unit. So, we divide the strokes in a line to sentence-strokes.

In summary, we have the following hierarchical grouping of strokes: strokes, sentence-strokes, and lines.

2.4 Layout

Acknowledgements

References

- BARNES, C., GOLDMAN, D. B., SHECHTMAN, E., AND FINKELSTEIN, A. 2010. Video tapestries with continuous temporal zoom. *ACM Transactions on Graphics (TOG)* 29, 4, 89.
- BORECZKY, J., GIRGENSOHN, A., GOLOVCHINSKY, G., AND UCHIHASHI, S. 2000. An interactive comic book presentation for exploring video. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, ACM, 185–192.
- CHI, P.-Y., LIU, J., LINDER, J., DONTCHEVA, M., LI, W., AND HARTMANN, B. 2013. Democut: generating concise instructional videos for physical demonstrations. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, ACM, 141–150.
- CHOUDARY, C., AND LIU, T. 2007. Summarization of visual content in instructional videos. *Multimedia, IEEE Transactions on* 9, 7, 1443–1455.
- CHUN, B.-K., RYU, D.-S., HWANG, W.-I., AND CHO, H.-G. 2006. An automated procedure for word balloon placement in cinema comics. In *Advances in Visual Computing*. Springer, 576–585.
- EKIN, A., TEKALP, A. M., AND MEHROTRA, R. 2003. Automatic soccer video analysis and summarization. *Image Processing, IEEE Transactions on* 12, 7, 796–807.
- HU, Y., KAUTZ, J., YU, Y., AND WANG, W. 2015. Speaker-following video subtitles. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 11, 2, 32.
- HWANG, W.-I., LEE, P.-J., CHUN, B.-K., RYU, D.-S., AND CHO, H.-G. 2006. Cinema comics: Cartoon generation from video stream. In *GRAPP*, 299–304.
- JACKSON, D., NICHOLSON, J., STOECKIGT, G., WROBEL, R., THIEME, A., AND OLIVIER, P. 2013. Panopticon: A parallel video overview system. In *proceedings of the 26th annual ACM symposium on User interface software and technology*, ACM, 123–130.
- KIM, J., GUO, P. J., CAI, C. J., LI, S.-W. D., GAJOS, K. Z., AND MILLER, R. C. 2014. Data-driven interaction techniques for improving navigation of educational videos. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, ACM, 563–572.
- KIM, J., NGUYEN, P. T., WEIR, S., GUO, P. J., MILLER, R. C., AND GAJOS, K. Z. 2014. Crowdsourcing step-by-step information extraction to enhance existing how-to videos. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, ACM, 4017–4026.
- KNUTH, D. E., AND PLASS, M. F. 1981. Breaking paragraphs into lines. *Software: Practice and Experience* 11, 11, 1119–1184.
- KURLANDER, D., SKELLY, T., AND SALESIN, D. 1996. Comic chat. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, ACM, 225–236.
- MONSERRAT, T.-J. K. P., ZHAO, S., MCGEE, K., AND PANDEY, A. V. 2013. Notevideo: Facilitating navigation of blackboard-style lecture videos. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1139–1148.
- MONSERRAT, T.-J. K. P., LI, Y., ZHAO, S., AND CAO, X. 2014. Live: an integrated interactive video-based learning environment. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, ACM, 3399–3402.
- PAVEL, A., HARTMANN, B., AND AGRAWALA, M. 2014. Video digests: a browsable, skimmable format for informational lecture videos. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, ACM, 573–582.
- RUBIN, S., BERTHOUSOZ, F., MYSORE, G. J., LI, W., AND AGRAWALA, M. 2013. Content-based tools for editing audio stories. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, ACM, 113–122.
- UCHIHASHI, S., FOOTE, J., GIRGENSOHN, A., AND BORECZKY, J. 1999. Video manga: generating semantically meaningful video summaries. In *Proceedings of the seventh ACM international conference on Multimedia (Part 1)*, ACM, 383–392.

Proving Trigonometric Formulas using Euler's Formula: $e^{ix} = \cos(x) + i \sin(x)$

$$\sin^2(x) + \cos^2(x) = 1$$

$$e^{ix} = \cos x + i \sin x$$

$$e^{-ix} = e^{i(-x)} = \cos(-x) + i \sin(-x) = \cos(x) - i \sin(x)$$

$$e^{ix} \cdot e^{-ix} = (\cos x + i \sin x)(\cos x - i \sin x)$$

$$1 = \cos^2 x - \cancel{i \cos x \sin x} + \cancel{i \sin x \cos x} - \cancel{i^2} \sin^2 x$$

$$1 = \cos^2(x) + \sin^2(x)$$

(a)

$\oint_C \vec{P} \cdot d\vec{r} = \iint_R (\nabla \cdot \vec{P}) dA$
 $d\vec{r} = dx \hat{i} + dy \hat{j}$
 $(\nabla \cdot \vec{P}) = \frac{\partial}{\partial x} P_x + \frac{\partial}{\partial y} P_y$
 $\oint_C P(x, y) dx = \int_a^b P(x, y_1(x)) dx - \int_a^b P(x, y_2(x)) dx$
 $= \int_a^b (P(x, y_1(x)) - P(x, y_2(x))) dx$
 $= - \int_a^b \left(P(x, y) \Big|_{y=y_2(x)}^{y=y_1(x)} \right) dx = - \int_a^b \left(\int_{y_2(x)}^{y_1(x)} \frac{\partial P}{\partial y} dy \right) dx$

(b)

Probability Lesson #22
www.actuarialpath.com

Uniform Distribution

$X \sim \text{Unif}(a, b)$

$(b-a)h = 1$
 $\Rightarrow h = \frac{1}{b-a}$

$f_X(x) = \frac{1}{b-a}, a \leq x \leq b$
 $F(x) = \begin{cases} 0, & \text{if } x < a \\ \frac{x-a}{b-a}, & \text{if } a \leq x \leq b \\ 1, & \text{if } x > b \end{cases}$

$E(X) = \frac{b+a}{2}; \text{Var}(X) = \frac{(b-a)^2}{12}$

(c)

f continuous on $[a, b]$

Fund. theorem of calculus

$$\frac{dF}{dx} = \frac{d}{dx} \int_a^x f(t) dt = f(x)$$

Every cont. f has an antiderivative $F(x)$

Connection between derivatives/integration

$F(x) = \int_a^x f(t) dt, \text{ where } x \text{ in } [a, b]$
 $\frac{d}{dx} \left(\int_a^x \cos^2 t dt \right) = \cos^2 x$
 $\frac{d}{dx} \left(\frac{1}{2} \ln(x^2 - 1) \right) = \frac{x}{x^2 - 1}$

(d)

Figure 3: Examples of lines (i.e. set of strokes that belong together semantically) output from our line-breaking algorithm. Our algorithm successfully identifies meaningful groups even from complex layouts with a mix of equations, figures and graphs.