La voix de synthèse comme alternative pour la vidéodescription (VD)
Synthetic voice as of of alternative for VD

Presented by
Ms Claude Chapdelaine
Senior Advisor
Vision et Imaging
CRIM

Montreal (Quebec)
November 15th, 2017
What is VideoDescription (VD)

A boy pretends to swim while lying on a chair.
Background on VD at CRIM

• Work done between
  • 2006-2007
    • Some of the actual VD elements can be detected automatically
  • 2008-2009
    • First production software and first player
  • 2010-2011
    • User testing with 43 visually impaired individuals across Quebec
  • 2012-2013
    • Software optimization and financing request
  • 2015-
    • Technical transfert in the industry in progress
    • Collaboration with BAnQ
    • Collaboration with ONF
Challenge and Automatic detections

• Challenge: Could image and sound detection help assisting a more efficient VD production?
  • To reduce production time
  • To lower production cost

• Proposed solution based on research and user experimentation:
  • Automatic detections integrated in the VDManager software
    • Shots detection (when scene and actors may change)
    • Key places for scene
    • Key faces for actors
    • Speech/Non-speech segments (where to include VD)

• Reduce rendering time
  • Integration of TTS engine for rendering of VD
Scene Detection

Principle of latent semantic analysis

Links for the grocery location.
Matches are within white boxes
Key Faces Detection

High level description of the methodology for the indexation of faces

Examples of clusters automatically generated
Production tool: VDManager
Advantages and constraints

• Advantages
  • Automatic detections provide
    • Cues and guidance to the scripters (information kept in a synchronized and interactive timeline)
    • An indexation of scene, characters and described actions
  • Rendering with TTS
    • Reduce significantly the production of the audio VD to be mix with the original sounds track

• Constraints
  • Finding an TTS engine that could :
    • Be integrated in the application
    • Can render many texts in batch mode
    • Offer good quality synthetic voices in the needed languages (Canadian French ...)
CRIM DVDPlayer
Usability Study on the Player

• In Situ study to better understand users’ needs with 43 participants across Quebec (collaboration of BAnQ)

• Production of an accessible Player that offered:
  • Contextual information
  • Two level of AD (standard and extended)
  • Recall functions (scene, actors and actions)
  • Native accessible navigation/feedback with good quality voice
CRIM Player features

- Contextual information giving with Information button
  - Synopsis
  - Actors (description, voice sample)
- Recall functions on current scene, actors and actions in that scene
- Advantage: recall functions are done in the production process
CRIM Player features

- Fully accessible via keyboard shortcuts and synthetic voice navigation
  - Getting to the starting menu
  - Feedback when using the player’s command
  - Bookmark feature included
- Compatible with screen readers (JAWS, Window-Eyes, ZoomText)
- In progress: development of Web and mobile versions
Rendering VD in synthetic voice

• Examples
  • Human vs TTS French
  • Human vs TTS English
Example : Human vs TTS French

ISABELLE AU BOIS DORMANT

Voix humaine
Example: Human vs TTS English

The Girl Who Hated Books

Human Voice
Rendering VD in synthetic voice

- Challenge
  - VD for film with multiple languages
    - Extended
    - Extended and voice over
CRIM Player - Rendering VD

- Allow selection of 2 different VD levels during playback
  - Standard => VD fits non-speech segments
  - Extended => DVD halts to render VD and restarts automatically
  - Extended with dubbing => DVD halts if necessary and render VD on voice over
Example : Extended VD without dubbing

Les Esquimaux de Netsilik

VD augmentée incluant la traduction (sans doublage)
Example: Extended VD mode with French dubbing

Comment les humains ont obtenu le feu

VD augmentée
et
doublage en français en mode augmentée
Example: Extended VD mode with French dubbing

Buffy

VD augmentée et doublage en français en mode augmentée
Future

• Changing voices in the same VD
  • Alternating between female and male voice if necessary
• Adapting production information for indexing film
• Migration of the player into a Web version and in mobility (App)
• Migration of production tool into a Web version

Merci - Claude.Chapdelaine@crim.ca

• Questions
• Suggestions