Localization on TV Applications

- A Position Paper for W3C Web and TV workshop –

Mahesh Kulkarni, Swapnil Belhe, Shubhanshu Gupta
Centre for Development of Advanced Computing (C-DAC)
Westend Centre III, Sector II, S/No. 169/1, Aundh, Pune, Maharashtra, India
{mdk, swapnilb, shubhanshug}@cdac.in

Introduction

In a diverse country like India with multi-lingual population separated by geographical distances and remote terrains having 22 official Indian Languages, localization plays a very important role to ensure the reach of any service to the grassroots.

The cable digitization and planned digital switchover of television broadcasting in India has opened new opportunities for services that enhanced viewing experience which offer much more than traditional TV. One such potential offering is localized interactive service. India is the 3rd largest television market. The total number of television households in India is 146 million with the TV penetration of 60 percent Because of its popularity and reach, TV is the largest medium for media delivery in India in terms of revenue [1]. In over 800 television channels in India, channels showing contents in local languages are more popular. Hence it becomes essential to bring the fruits of internet to the large number of people in local languages.

In order to tap the huge potential of TV in Indian context, we created television based learning framework. The application discussed in this paper is developed to reinforce distance learning initiatives through television medium making Television a 'Virtual University' by bringing knowledge to all in local languages. The application provides value added interactivity that keeps the learner engaged to the educational programme, this promotes viewers participation resulting in an immersive and lasting learning experience. The application is designed to reuse e-learning contents for web. This offers the learner-viewer with the facility to take the lecture notes, self evaluation tests and access supplementary information, making learning a fun as well as a holistic experience in their local language.

The current standards such as HTML5 and CSS provide good support to render international languages. However, there are some issues related to Indian language layouts. These issues will be extended to TV medium.

Challenges

The current CSS standard has following issues with respect to Indian languages which would become equally important to resolve for issuing T-learning contents on television medium.

- Styling of first letter pseudo element in Indian languages
- Vertical alignment of Indian languages
- Underlining of some Indian languages
- Inter character spacing
- Word (hyphenation) rules
- Line breaking

W3C's Multimodal Interaction Working Group (MMI-WG) is designing various multi-modal interfaces for different digital mediums including Television. In order to get interactive experience during television based learning, one should be able to communicate with the TV in local languages. This can be achieved by means like voice, text, gesture, handwritten etc. Indian languages contain many characters hence it become difficult to accommodate them in limited space on keyboards. In such cases ability to use multimodal interfaces apart from textual becomes necessary.

Current Research

The ongoing research work is to enable localization support to set-top-boxes in terms of Indian language fonts, layout engines, limited key Inputting schemes, and multilingual component library for TV applications and Indian Sign Language for broadcast.

1. Inputting Mechanisms

Inputting mechanisms enables typing in Indian Languages on set-top box through limited keys of remote control. The following schemes have been devised for Indic inputting.

a. Two-Key Inputting Mechanism:

This mechanism is based on C-DAC Two key Inputting Mechanism. In this system each key from (0-9) is mapped to 5 characters of a particular language. Hence for typing a single character, multiple tabbing is required. When we press a key; all character combinations will be shown in suggestion box, '# key' is used as switching key for second layer characters. Developers can use "GSinglelineEntry" and "GMultilineEntry" components of this system in their applications to enable inputting in Indian languages.

There are two ways of typing using this tool:

a. Using multiple tabbing: Numeric key 1 is mapped to these five characters ['क','ख','ग','घ','ङ']. Hence for typing 'ग', we need to press numeric key three times.

b. Selecting character from Suggestions: When we press any numeric key of TV remote, all characters mapped to that key will be displayed as suggestions. For typing we need to select from suggestions using 'OK' key of remote.

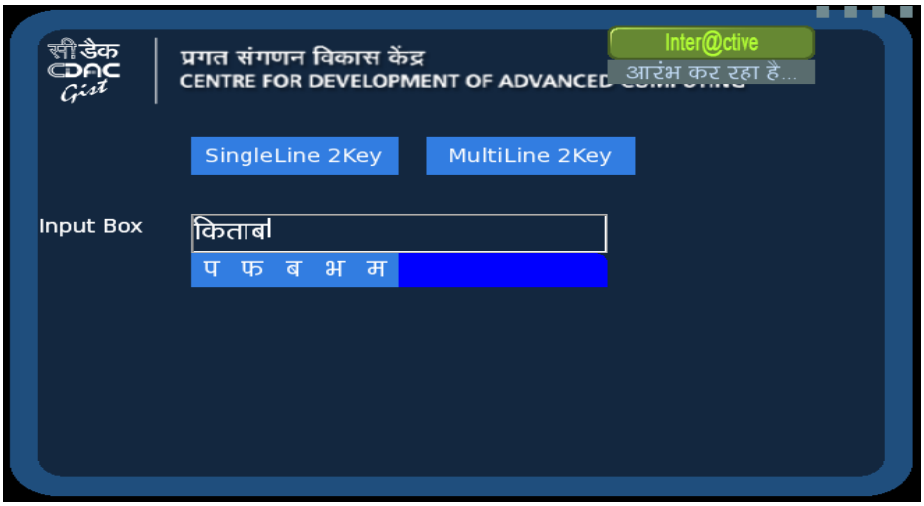
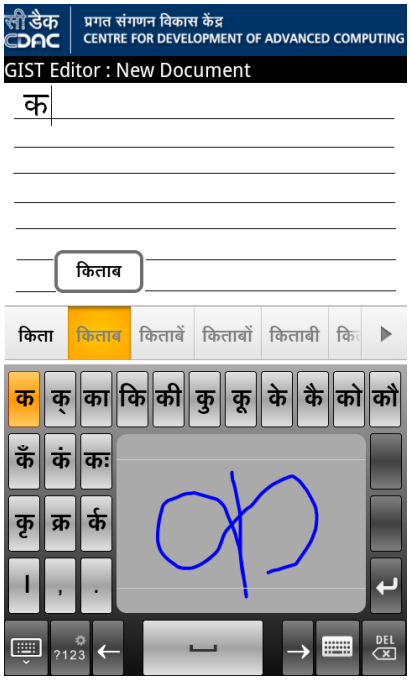


Figure 1: Two-Key Inputting Mechanism on Set-top-box Simulator



GIST On-Screen-Keyboard:

This mechanism for typing in Indian language on Set-top-boxes presently supports English, Hindi, Gujarati, Bengali, Kannada, Malayalam, Marathi, Punjabi, Tamil and Telugu languages with Unicode 6.2 enhanced INSCRIPT layouts. This is being developed as a component “OSKPanel” for use by third party application developers.

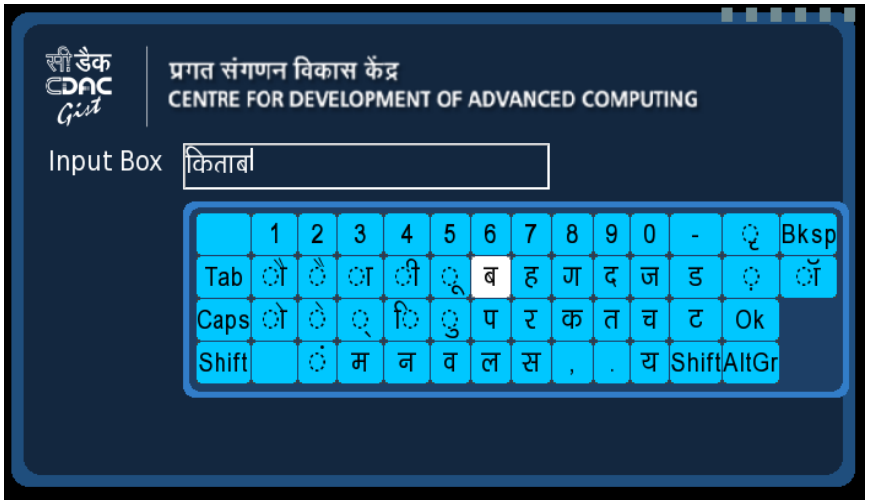


Figure 2: On-Screen-Keyboard on Set-top-box Simulator

2. Font Rendering and lay outing technologies

Support of various fonts and inputting mechanism for the TV platform

3. Indian Sign Language

- Standardization efforts for pushing Text to sign language on to the TV medium
- Creation of captioning system customized for Indian scenario.
- Creation of News / Contents in Indian Sign Language (ISL) for broadcasters.
- Creation of teaching and learning materials for deaf persons.
- Guidelines generation for captioning multimedia contents for broadcasters, production industry etc.
- Rendering of different Awatars (synthetic actors) on the TV medium

4. T-Learning (Television based learning)

Television plays a vital role in 148 million [1] Indian homes as a medium for infotainment, edutainment and entertainment. It is also a source of Educational TV (ETV) programs available as a public service to nurture India

Govt. of India has a major focus on the "Technology Enhanced Learning" in its 12th Five Year Plan (2012-2017) as it provides universal access to information and services to the public in their own language

Need for interactive (bi-directional) TV learning service:

Learning process always involves an active mind, demands learner’s attention and curiosity but Television has always been perceived as a passive medium. Modern digital television offering interactivity rich multimedia experience answers to this essential problem of delivering

Distance Learning through Television medium. It encourages viewers to more and more interact with the TV to participate in the active learning environment that too in their local language. Dissemination of knowledge in local languages is an important requirement for a linguistically and geographically diverse country like India.

Challenges

1. Set-top-box Localization
    a. Many set-top-boxes uses older fonts which may not contains few newly added characters and hence they require an update.
2. Limited inputting keys
3. Accessibility
    a. Different Indian languages occupy different screen area hence a scheme to logically organize the text is required.
4. Limited memory footprint
5. Limited memory for Indic Fonts
6. Addressing the hard-of-hearing

Conclusion

Support for localization components/API's is required .

- *Localization as a business opportunity:* The industry has to recognize that users with impairments are market-relevant user groups, generating corresponding revenue when enabling access to services and products. This is especially true for the growing group of elderly citizens.
- *Joint efforts:* All stakeholders from industry, research and public authorities have to collaborate effectively to remove the current barriers in development of localized services.
- *Guidelines:* Currently there are already many guidelines and best practices available, offered by most of the standardization bodies.
- *Dissemination & education:*.
- *Localized user interfaces:* .
- *Open Source:* Localization software frameworks have to be open source and freely available, to ensure wide uptake and efficient integration by the industry.

References

1. CDAC Two Key Mechanism Layout.