An Architecture for Real Time Television Audience Measurement

Debnath Mukherjee, Tanushyam Chattopadhyay, Siddharth Bhattacharya, Avik Ghose, Prateep Misra
{debnath.mukherjee, t.chattopadhyay, siddharth.bhattacharya, avik.ghose, prateep.misra}@tcs.com
TCS Innovation Labs, Tata Consultancy Services
Bengal Intelligent Park, Sector V, Salt Lake Electronics Complex, Kolkata – 700091, India

Abstract- **Currently, audience measurement reports of television programs are only available after a significant period of time, for example as a daily report. This paper proposes an architecture for real time measurement of television audience. Real time measurement can give channel owners and advertisers important information that can positively impact their business.**

We show that television viewership can be captured by set top box devices which detect the channel logo and transmit the viewership data to a server over internet. The server processes the viewership data and displays it in real time on a web-based dashboard. In addition, it has facility to display charts of hourly and location-wise viewership trends and online TRP (Television Rating Points) reports. The server infrastructure consists of in-memory database, reporting and charting libraries and J2EE based application server.

I. Introduction

Television Rating Points (TRP) is a measure of the popularity of a television channel or program. TRP is expressed as a number between 1 and 100 where one rating point equals 1% of a target television audience population. TRPs are measured using statistical sampling techniques from a target population.

TRP ratings are important for television broadcasters, media agencies, advertising agencies and advertisers. It has a huge impact on advertising expenditure and TV program scheduling. Hundreds of billions of dollars worth of air time buying and selling is done each year by media houses and companies based on TRP ratings.

Current methods of automated TRP measurements is based on placement of devices known as “People Meters” in the homes of the sample population. These devices are connected devices that measure viewing habits and periodically send reports to a back end system. Current methods used by “People Meters” are explained as follows.

1) **Audio Matching** – In this case the People Meter captures the audio content of a TV program , compresses the data and sends the data to a central backend server. In the server, the audio samples are compared with stored program audio data and thus viewership data is measured. This method is complex to implement and requires substantial investments in backend systems.
2) **Frequency Measurement** – This works with analog transmission systems, where the People Meter tracks the frequency of the tuned TV channel and sends this information to the backend. The backend maps the captured frequency with the channels and programs and is thus able to track viewer behavior. This method is the most popular method used in the country but is also most susceptible to errors / inaccuracies as there is no regulation or monitoring on local cable service providers with respect to the frequency of transmission.
3) **Watermarking** – In this technique watermarks are inserted in the program feed at the broadcasters end and is detected by the People Meter. The people Meters then transmits the detected watermark and timestamp to the backend. This method suffers from the fact that each program transmitted must be watermarked and that active cooperation from all broadcasters is required.
4) **Visual recognition** – In this case the People Meters looks for visual patterns and images from the displayed screen and thus tries to ascertain the program being watched. The proposed system uses this method of. TRP measurement.

The current TRP reports are sent on daily, weekly, monthly or quarterly basis, not in real time. Real time TRP reports can give an indication of current popularity of channels and thus help advertisers place their ads effectively.

The proposed system elegantly solves the problem of real time TV viewership tracking by using channel logo detection techniques. It uses sophisticated image processing techniques to detect the channels in real time based on stored signature of channel logos. The image processing is done by Digital Signal Processors present in the set top box known as HIP (Home Infotainment Platform). The detected channel and location is sent to a backend analytics server which shows the viewership data through dashboards and online reports. In some DTH systems, the channel being viewed can be found, however the set top box allows processing of analog RF cable as well. DTH has lower reach than cable, so this is significant.

The unique contribution of this paper is that it presents a novel application which helps advertisers and broadcasters view current trends of television viewership both in graphical and report format. It also reports current (real time) TRP of the various channels. The real time TRP will enable advertisers place advertisements based on current popularity of channels and also based on location factors.

III. RELATED WORK

The related work is in the area of "digital signage" which Wikipedia describes as "a form of electronic display that shows information, advertising and other messages." These signs are displayed in out-of-home locations such as retail stores and corporate buildings. Some software such as Trumedia [6] helps advertisers give targeted advertisements by e.g. analyzing the current audience based on video analytics on the faces of the audience captured by a camera. The data about the audience is sent back to a backend server which helps display reports about the current audience. In TruMedia's integration with Cisco Digital Media Player, the real time information about the audience can be used to select the next message to play, thus enabling targeted advertising. Quividi [7] and CognoVision [8] are two other players in this market.

Our solution is for in-home TV displays and helps measure channel popularity in real time. This channel popularity can be used by advertisers to place ads at the right time.

III. PROPOSED SYSTEM

The TV broadcast is routed to the device via composite A/V IN and output is given to the television via composite A/V OUT. Logo detection algorithms are run on the video using the digital signal processor in the HIP box. On recognition of the channel being viewed, the box constructs a tuple (channel identifier, location, timestamp) and sends it to the analytics server which displays various charts and reports.

The benefits of TRP tracking using HIP is as follows:

1) **Low cost TRP tracking** - The HIP platform is an extremely low cost device. It costs approximately an order of magnitude less than existing People Meters. Most importantly, TRP tracking is only one of the many applications that the device supports.

2) **Ease of deployment** - This is a software only solution that sits inside HIP. The primary purpose of HIP is not audience data measurement. This is an add-on software functionality.

3) **Real time audience measurement** – Data about viewership will be available in real time. This is as opposed to end of day or end of week data that current systems now provide.

4) **Increased accuracy** – Current systems are based on frequency monitoring are considerably less accurate since frequency of transmission in cable TV networks are changed arbitrarily from time to time by local cable TV operator.

5) **Larger sample size** – HIP being widely deployed will provide much larger sample sizes than the existing Peoplemeter devices.

6) **No possibility of data tampering by vested parties** – Since data is available to subscribers in real time, there is no scope for any manipulation of any stored data or doctoring of reports. Also secure channels from HIP platform will ensure complete integrity of measured data.

The deployment of the proposed solution is shown in Fig. 1.

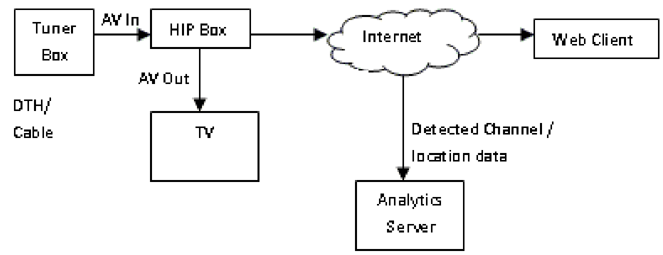

Fig 1. Deployment of the proposed solution

The basic working principle of this measurement system is as follows:

- HIP set top boxes detects logos of channels being viewed when channel is changed by the viewer
- HIP box transmits the detected channel id, current time stamp and IP address of the HIP box to an internet hosted analytics server using secure channels
- Geographic location is obtained using IP address. Each HIP box has a (static / dynamic) public IP address from which the geographic location of the box can be determined from Google API.
- Electronic Program Guide is obtained from internet sources
- Analytics server analyzes the data and creates dashboards, mashups, heat maps, real time trending and charts.
- Data from analytics server is available to authorized clients using a internet connection and browser

IV. CHANNEL LOGO DETECTION

A. Overview

The overview of channel logo recognition can be found from [1]-[3]. The steps are described in Fig. 2.

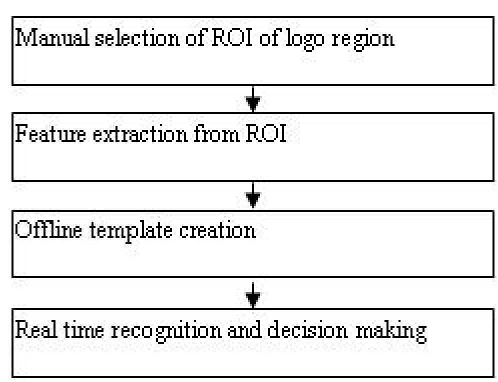

Fig. 2. Overview of Channel logo recognition

B. Manual selection of Region of Interest (ROI) of logo region

Initially the videos of all channels are recorded to create a single video file. Manual annotation is performed on the video file to generate an xml file containing channel code, start frame number and end frame number.

This video is played using a tool that enables the user to select the ROI (logo region) from the video using mouse. The tool takes the annotated xml file as input to generate a output xml file with ROI coordinates, height and width of ROI and features of the ROI.

The feature consists of quantized colors of the pixels in the ROI. We have used quantized colors to reduce the size of the template file.

C. Method of color quantization

The input candidate channel logo region image is UYVY format, which is converted to HSV color space to have separate information of color, intensity and impurities. These HSV values are quantized into 36 bins as described in [3].

D. Recognition of Channel logo

Once the candidate logo regions are found, each pixel is quantized to 36 values. Then the features of each candidate region are matched with the features of template logos stored in the database. In the proposed method a fuzzy multifactorial approach is used to recognize the channel logo. The factors and the reason behind selecting these factors are as below. (Details of the logo detection method can be found from [3]).

Bhattacharya distance: We have used a factor (f_bhat) based on that feature to find the histogram similarity between the candidate and the template. Details about this distance can be found from [5].

Crossing count and run length similarity: These are two shape invariant features used in OCR. We have constructed a Similarity matrix of size 36x36 and find the similarity between template and candidate using correlation coefficient. As the normalized correlation value also lies in the interval of (0,1), the factor value also in the same interval. The method of construction of the matrix is as below:

$$i \leftarrow quant[y][x],\ j \leftarrow quant[y][x+1],$$
$$cross\_mat[i][j] = cross\_mat[i][j] + 1,$$
$$\forall\ y \in (0, height)\ \forall x \in (0, width-1)$$

Where quant is the array of pixels with quantized color value, cross_mat is the matrix storing the crossing count feature

Flusser and Suk image moments: One simple Affine Moment Invariants (AMI) used for channel logo recognition. In [4] authors have used four such features. But other than the first one, rest is computationally expensive for implementation of a real time system in a DSP platform and it is found experimentally that the Affine Moment Invariant 1st Order Transformation (AMI 1) itself produces very good result.

Aspect Ratio: We have used another factor (f_asp) based on this feature can distinguish between different families of channel logos very efficiently. It is found that the aspect ratio of the logos of Star family varies with the same of Zee family. The method of assigning fuzzy weight to each of those factors is described in [3].

Construction of evaluation matrix

The evaluation matrix is formed using the factors values as described below

$$V = \begin{pmatrix} f\_bhat_1 & f\_bhat_2 & \ldots & f\_bhat_n \\ \cdots & \cdots & \ldots & \cdots \\ \ldots & \ldots & \ldots & \ldots \\ f\_asp_1 & f\_asp_2 & \ldots & f\_asp_n \end{pmatrix}$$

Construction of Additive Standard Multifactorial (ASM) function

Now, from the evaluation matrix V it is difficult to obtain any solution of the decision-making problem. So we define a mapping function M_m that maps the m-dimensional vector $\mathrm{f} = (\mathrm{f}_1, \mathrm{f}_2, \ldots.., \mathrm{f}_\mathrm{m})$ into a one dimensional scalar i.e. $M_m(f) = \mathrm{M_m}\ (\mathrm{f}_1, \mathrm{f}_2, \ldots.., \mathrm{f}_\mathrm{m})$

we apply ASM on V to obtain the multifactorial evaluation

$$\mathrm{V}' = (\mathrm{v}_1, \mathrm{v}_2, \ldots\ldots\ldots\ldots\ldots\ldots.., \mathrm{v}_\mathrm{n})$$

Where $\mathrm{v_i} = \mathrm{M_m}(\mathrm{v_{1i}}, \mathrm{v_{2i}}, \ldots\ldots.., \mathrm{v_{ni}})\ \forall\, \mathrm{i} = [1, \mathrm{n}]$.

Now we define a mapping function $\mathrm{M_m}$ as a simple arithmetic average over the m number of factors and thus it is defined as

$$M_n(d) = \frac{1}{m}\sum_{i=1}^{m} d_{in}$$

So we finally obtain the decision making matrix V' which is a row matrix and each element represents the confidence score of membership of the candidate region matching with the i^{th} logo.

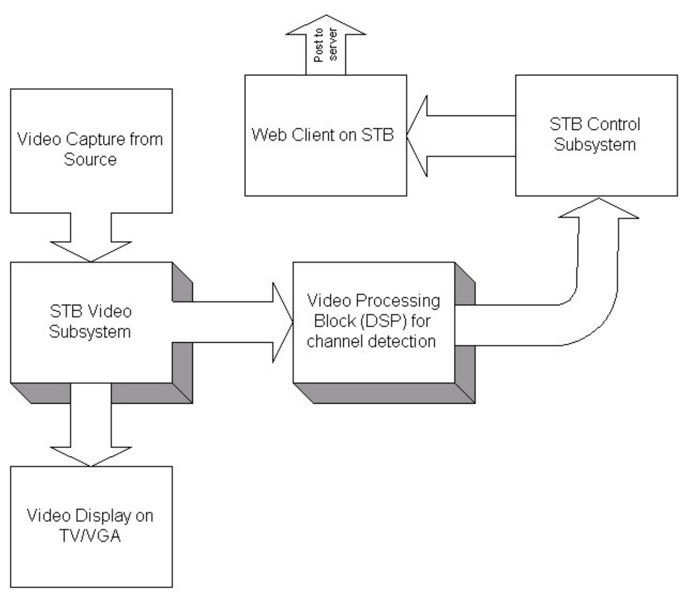

Fig 3: Architecture for sending data from the set top box to the analytics server

V. Reporting Of Data To The Analytics Server

The overall enhancements to the set top box is depicted in the architecture diagram (Figure 3). The sequence of events is as follows:

1. Video is captured using set-top-box (STB) video capture block
2. This video is then processed by logo detection algorithm running on DSP
3. The algorithm returns a channel id which is passed to the control sub-system of STB
4. The channel id and the time is recorded and fed to the box HTTP client
5. The HTTP client then posts this data at a 30 second (configurable) timer to the server

The web-poster utility runs in a timer loop and keeps posting channels watched, serial number (unique id), timestamp, location id and current time.

VI. Data Analytics On Streamed Data

A. *Overview of the Analytics Solution*

The architecture of analytics is shown in the following diagram.

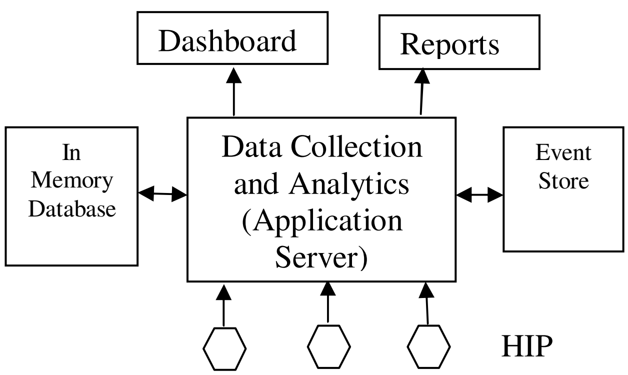

Fig. 4. Analytics Architecture

The steps in the data processing are explained in the following paragraphs.

Data is sent from the HIP boxes to a data collection application server implemented using J2EE technology. The data tuple consists of the location, timestamp and channel identifier.

The data is stored in an in-memory database. The database stores only a few minutes' worth of data. Data older than that is deleted. This database serves data to a real time display module which periodically takes snapshots of the data and displays on the web. Since real time current data is displayed, old data can be deleted; which helps to keep the size of the table small and improves the performance of the real time display module. The data in the in-memory database is also used to show the viewership statistics at various locations on a map of the country. This is done using Google Map Application Programming Interfaces.

The data is also stored in an event store, a relational database management system, for online reporting of:

- Television Rating Points (TRP) for a selected channel for a selected time period.
- Hourly trends of channel performance
- Location-wise trends of channel performance

Reports are published both as PDF and as HTML.

TRP for a channel during a time period is estimated as:
(total viewed time) / (total available time) * 100
Where "total viewed time" is the sum of the viewership of a channel during the selected time period and the "total available time" is the duration of the time period multiplied by the number of set top boxes deployed.

The system has been built using open standards such as Java/J2EE and SQL. This allows the system components to be replaceable with products from different vendors if needed.

The technology used in the implementation is as mentioned in Table I below.

TABLE I
TECHNOLOGY MATRIX

COMPONENT	TECHNOLOGY
Application Server	JBoss
Reporting library	Jasper Reports
Relational Database	MySQL InnoDB Engine
In-Memory Database	MySQL Memory Engine
Geographical Mashup	Google Maps API

B. *Features of the analytics solution*

The web frontend implements visualization of data in several "views". The available views are:

All Channels View – This view shows the performance of all the channels as bar charts – with viewership on the Y axis and the various channels on the X-axis.

Single Channel View – This view allows the user to select a location and a channel and display the viewership. The channel is shown on the X-axis and the viewership on the Y-axis.

City-wise Channel View- This view shows the performance of channels being watched at a particular city as bar charts - with viewership on Y axis and the various channels on the X axis.

Country-wide View - This view shows the performance of a particular channel across all geographic locations. Markers are placed at various locations on the map and on clicking on the marker a pop-up shows the viewership.

Reports View – This is divided into 3 kinds of reports

1. **Hourly Trend Report**- This view shows the viewership trend chart for a particular channel as it varies with time in the form of a line chart with number of viewers on the Y-axis and Hours on the X-axis. The data points are also shown in the report (hour, number of viewers).

2. **TRP Viewership Report**- This view shows the TRP value for a particular channel in a specified time window.

3. **Location based Trend Report**- This view shows the viewership trend chart for a particular channel and a specified location as it varies with time in the form of a line chart with number of viewers on the Y-axis and hours on the X-axis. The data points are also shown.

C. *Modules of the application*

The modules of the solution are explained below:

Data Collection Module – This module implements RESTful Web Service [9] to receive the data posted by the HIP boxes. A servlet processes the requests and writes to the in-memory database and the event store.

Image Generation Module – The real time display is implemented by taking snapshots of the data in the in-memory database periodically and generating PNG files. Different PNG files are generated for different views. Each image shows the viewership on the Y axis and the channels on the X-axis. The image generation is a background process.

Database Cleaner Module – As mentioned before, the old data in the in-memory database has to be deleted as old data is not required for showing real time charts. To achieve this, a background process is implemented.

Reports Module – Reports are generated using the JasperReports library. Each report shows a plot of the data at the top of the report and the actual data points below it in a tabular format.

Geographical Mashup module – The geographical plot is implemented using Google Maps API. Markers are placed at various locations where television audience measurement is done. On clicking on a marker, the viewership at that location is shown. This is a real time chart and Ajax technology is used from the web page to fetch the real time data from the in memory database.

VII. RESULTS

A. *Results for Channel logo detection*

The PC version of the code was tested over 280 different recorded videos from different channels. However, the current DSP based system can support upto 13 channels with recall and precision rate almost equal to 1. Recall (r) and Precision (p) can be defined as:

$$r = \frac{c}{c+m}, \; p = \frac{c}{c+fp}$$

where c is the number of correct recognition, m is the number of misses and fp is the number of false positives. For each channel we have tested the algorithm for 1 hour TV sequence taken at 20 different instant of time. The current software supports 13 channels. The average time to detect the logos is 1.87 milliseconds with a maximum of 1.95 ms and minimum of 1.8 ms. Code is developed using Texas Instruments(TI) Codec Engine framework.

B. *Results for Analytics*

This section shows screen shots of our TV Audience Measurement application. Figure 5 shows the geographical mashup. Figure 6 shows the hourly trend report. Figure 7 shows the TRP report and Figure 8 shows the All Channels view.

We simulated the set top boxes as threads running in a Java program. The various timing configurations are as follows:

1. The web poster on the STB posts once every 30 seconds
2. The image generator runs once every 5 seconds
3. The web screens are refreshed once every 10 seconds

Additionally, the network delay to send data from simulator to server was found to be 2 seconds.

Measurement on an Intel dual core 2.33 GHz processor with 2 GB RAM showed the average latency to reflect a channel change as 47.6 seconds.

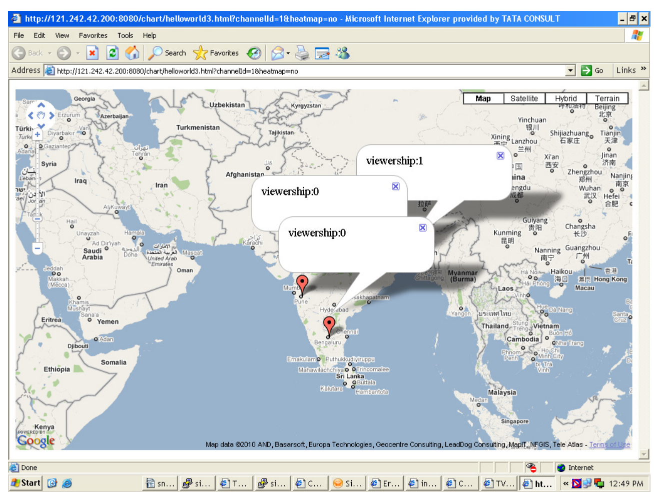

Fig. 5 Geographical mashup

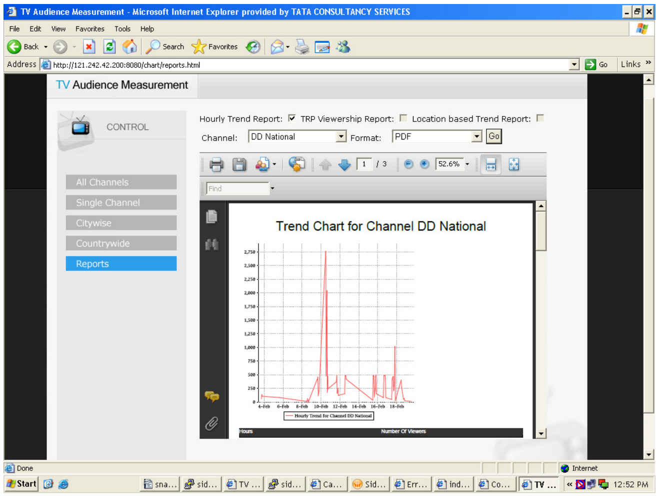

Fig. 6. Hourly trend report

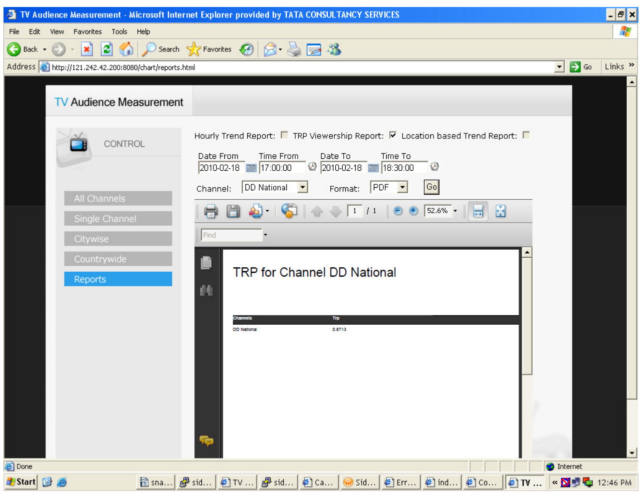

Fig. 7. TRP Report

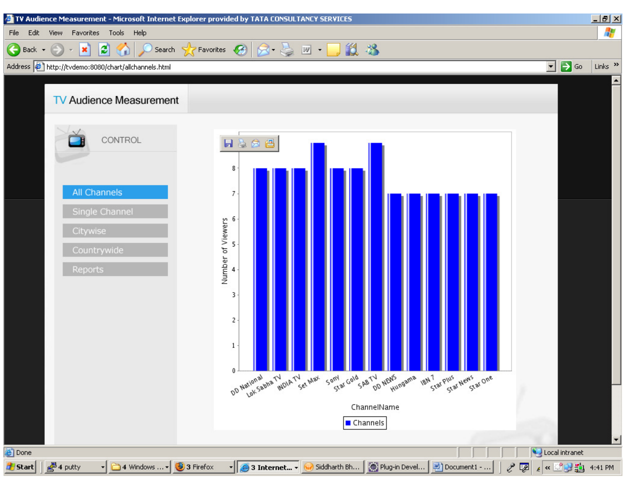

Fig. 8. All Channels View

VIII. CONCLUSION AND FUTURE WORK

In this paper we have presented a novel application for real time measurement of TV audience. The application detects the logo of the currently playing channel and reports the channel and location to an analytics server. The analytics server helps the user to discover trends in viewership across channels and locations. This application would be useful for TV broadcasters and advertisers in their business efforts. In future work, we plan to investigate the possible application of Bayeux protocol [10].

IX. ACKNOWLEDGMENT

We acknowledge the involvement of Computational Research Laboratories (CRL) in this project.

REFERENCES

[1] T. Chattopadhyay, Aniruddha Sinha, Arpan Pal, Debabrata Pradhan and Soumali Roy Chowdhury, "Recognition of Channel Logos from Streamed Videos for Value Added Services in Connected TV,"Proc. of the 29th International Conference on Consumer Electronics (ICCE'11), Jan, US, 2011

[2] T. Chattopadhyay, and ChandrasekharAgnuru, "Generation of Electronic Program Guide for RF fed TV Channels by Recognizing the Channel Logo using Fuzzy Multifactor Analysis", Proc. of the 14th International Symposium on Consumer Electronics (ISCE'10), 7-10 June, Germany, 2010.

[3] T. Chattopadhyay, Sounak Dey, Pritha Bhattacharya, Arpan Pal, "Electronic Program Guide for RF fed TV Channels," Proc. of the UKSim 4th European Modelling Symposium on Mathematical Modelling and Computer Simulation, (EMS 2010), Page(s), , Pisa, Itally, 2010

[4] J. Flusser, and T. Suk, "Affine Moment Invariants: A new tool for character recognition," Pattern Recognition Letters, Vol 15, page(s) 433-436, 1994

[5] T. Chattopadhyay, P Biswas, B. Saha and A. Pal, "Gesture Based English Character Recognition for Human Machine Interaction in Interactive Set top box Using Multi factor analysis", 6th Indian Conf. on Computer Vision, Graphics and Image Processing, ICVGIP, Page(s) 134-141, India, 2008

[6] Trumedia, http://http://trumedia.co.il/

[7] Quividi, http://www.quividi.com/

[8] Cognovision, http://www.cognovision.com/

[9] Representational State Transfer, http://en.wikipedia.org/wiki/Representational_State_Transfer

[10] The Bayeux Specification, http://svn.cometd.com/trunk/bayeux/bayeux.html