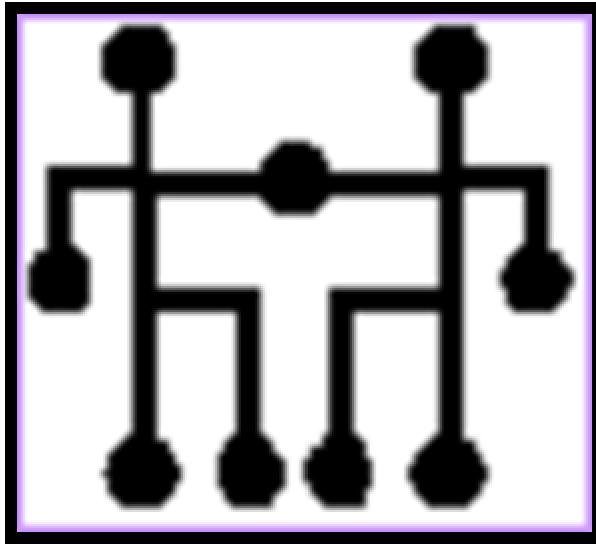Howard University
Department of Electrical and Computer
Engineering

Radio Television Instant Search Engine

Fall 2011

The Cyberplex Blueprint Team:
Christopher Caesar
Jonathan Murphy
Bryant Beeler

_____
Advisor Name   Signature   Date

_____
Class Instructor Name   Signature
        Date

Table of Contents

Introduction

Objective

The objective of this proposal is to outline the details of the design proposed by the Starplex Blueprint Team for the challenge set by DigiClips, Inc. The project will provide a web based search engine to capture stream and analyze television, radio, social networks (Facebook, Twitter, LinkedIn, YouTube, etc), on-line newspapers, magazines, blogs, etc. for general public and public relations executives for personal and organizational use for media analysis. The functionality is to provide through a key word search through close caption to match and parse the close caption to still pictures, click and play the video/audio of the segment searched for and generate text reports.

Background

DigiClips was founded in 2003 as a Colorado C media monitoring corporation, which monitors all types of media from newspaper websites magazines blogs and news correspondent sites in 210 US markets and 89 International Markets. Services includes the research, compilation and distribution of news stories and commercial tracking in multiple formats: DVD, CD, audio clip, transcript and log summaries delivered by mail, email or fax.

Some of the clients the DigiClips service are media relations personnel from universities, government, and various industries. The primary goal of DigiClips, Inc. is to deliver media content to our clients within minutes of the content being aired through different media delivery methods with the primary method being via the Internet. By delivering within minutes of air time over the internet, DigiClips, Inc. will create a self-service method for clients to achieve timely planning and responses to public relations (PR) concerns.

Problem

Problem Definition

Our goal is to create a real time web search engine capable of analyzing media outlets for PR purposes.

Assumptions

We want to be able to design a search engine sophisticated enough to give accurate results, but be simple and easy enough for even the computer illiterate user to use. This search engine will be able to access different types of media such as video, radio, and social networks.

Design and Measurement Requirements:

Our search engine application will be written in C++ software language. SQL is also to be used in order to capture the closed caption with the video/audio data and import it into a timeline into

the SQL relational database. The front end web site is to be developed for data to be accessible to users. A PCI express slot is also required to allow connection for the hardware (namely the HVR board). Access to cable television is also required in order to be able to pull a video signal. There aren't any measurement requirements of note to speak of, only preference of the storage size for the database

Overall Function:

Providing a key word search through close caption to match through close caption to match the closed caption to still pictures, click and play the video/audio of the segment searched for and generate text reports.

Performance

- User should be able to receive search results either by doing a manual search or by receiving alerts to a profile account they have setup with DigiClips

- Search results will be based on variables put in by the user

- Range of results can also be related to a time interval put in by the user.

- A raid controller with fifteen 1.5 TB hard drives will serve as the main database.

Guidelines

We will be working along with Mr. Robert Shapiro of DigiClips Inc. Technical support will be available though an IT administrator who also works with DigiClips. It is to be noted that there are no developmental restrictions, only suggestions and guidelines based on what has been done already. Our ability to communicate with DigiClips is reliant on exchanging emails with Mr. Shapiro who is an owner of the company. There is also a computer called digiclipstestbox where we would be able to log into and do our work on the project remotely, a must given the distance between the team and the Colorado based company DigiClips.

Required Compliance

All software we will be using will be open-source, so there are no copyright legal issues. However we did have to sign a confidentiality agreement before being allowed access to the development software. This prohibits us from using the software or the project in general for financial gain, or to aid any potential competitors with DigiClips. It allow prevents us from disclosing any information about DigiClips client list, trade secrets, or any other company information without authorization.

Current Status of Art

Similar Product in the market

There are currently similar components out on the market but there is not a similar product out on the market that is capable of meeting all the specs that were laid out by Digiclips. Digiclips has taken the steps toward developing the product but has been unsuccessful in the past. There is no similar product that works successfully, but the team will use past attempts and research to put together a product that solves the problem and meets all the specs. This will be the first time that all of the components are put together in a format and interface that will solve the problem specific to the problem given to the team.

Similar Problem Already Solved?

The problem has already been solved to some degree. There are already programs out that will allow for closed captions and speech to be dictated and the technology for the other tasks has been out for some time now. The challenge is getting the different components to communicate and intergrate the components while maintain a user-frienly interface. The problem has already been solved on the level of the individual components. The problem has not been solved as far as creating a program and making the different components interfaceable.

Available Technology

There is new technology available in comparison to the products that are already in place. There have been advances in microprocessors. There are similar products but the group will be using open source material from existing products like:

Google- Search the world's information, including web pages, images, videos and more. Google has many special features to help you find exactly what you're looking for.

Bing-Bing is a search engine that finds and organizes the answers you need so you can make faster, more informed decisions.

Yahoo-The search engine that helps you find exactly what you're looking for. Find the most relevant information, video, images, and answers from all across the Web.

Weakness of the Current Solutions

The weakness of the current state could best be described as repetitive. The difficulty with the search engines today is that they depend on other sites to retrieve the information they provide for their own sites. In addition, so when something happens it isn't instantly available to one person but once one web search engine has access to it, they all have access to it.

Approach for Solution

Primary Solution Design

Our goal is to create the latest most innovative and captivating instant radio provider. The search engine will work mainly off of the closed captioning feature where the user will explain exactly what is being requested. The search engine will then sift through the database for the exact words retrieved from the user. In addition, to the web content it will also retrieve scripts of video clips to also provide rich media to the user which will be collected by the integrated Hauppauge HVR 2250 board. Furthermore, team Cyberplex Blueprint will also renovate the DigiClipsinc.com website in an attempt to make the web service more edgy and user friendly.
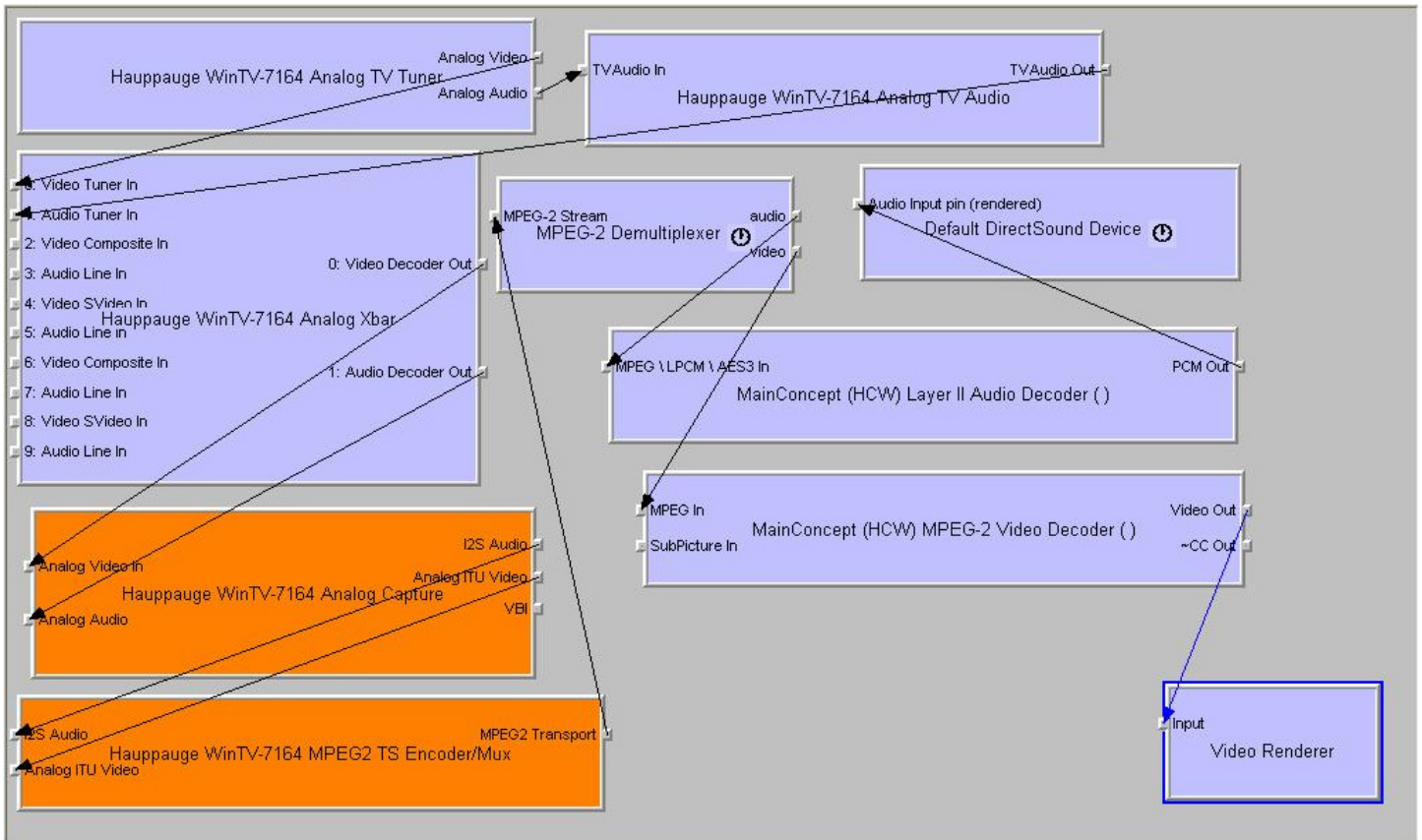
Concept Study

Understanding the aforementioned requirements, guidelines, and restrictions for the RTV search engine, and gain a basic understanding of the work and research that has been done in the past in order to apply it accordingly to the immediate task.

Concept Development

Brainstorming, and then developing ideas for a functioning program that will meet all the specifications and also be user-friendly and present ease of use to the most computer illiterate user.
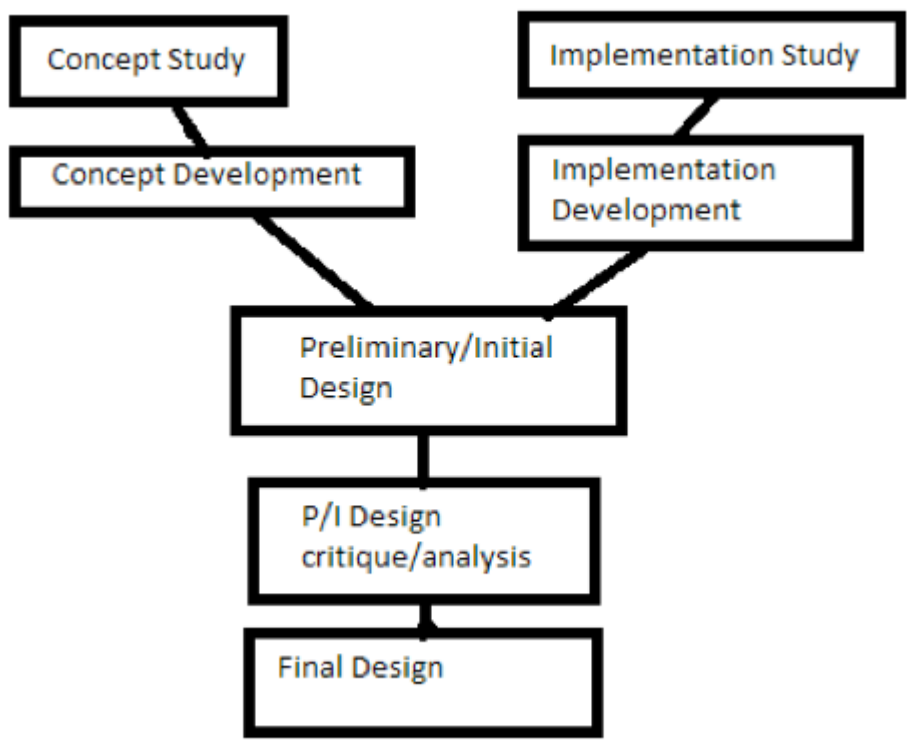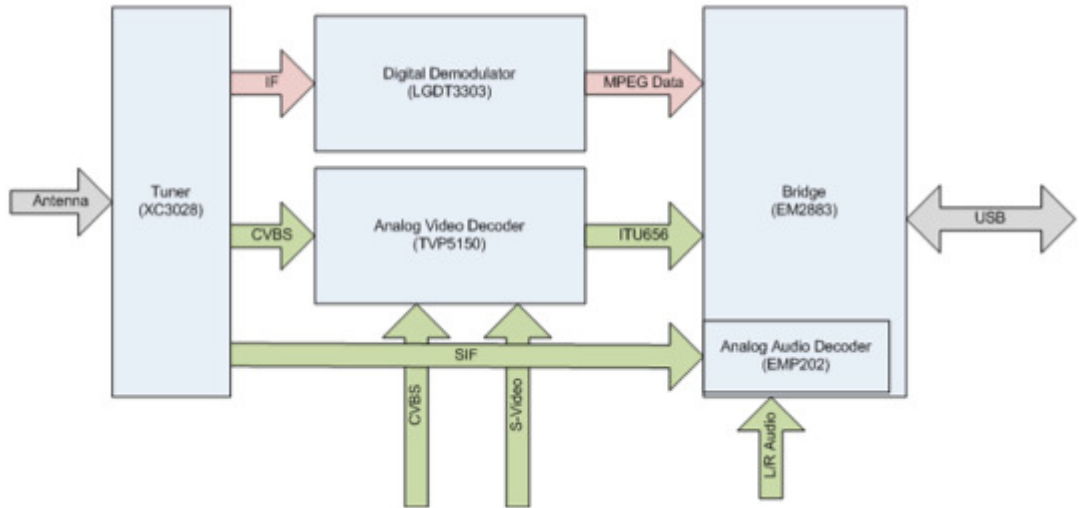
Implementation Study

Understanding and deciding what components needed to create a basic functional search engine will be the most fundamental step in our research and planning stage; if we understand what components are absolutely necessary we can then begin going our own way from those footprints already made. In addition, coupling the implementation of the software with the hardware will the cornerstone to the completion of the project for the hardware will provide a pertinent aspect to the overall project.

Implementation Development

Obtaining the components necessary for complete assembly of the hardware required to put together the RTV search engine. Also, obtain the extra programming and database knowledge necessary to createe the program, interface, and location to store the data. The team will also have to constantly obtain knowledge and suggestions from Digiclips in relation to the progress of the team and any changes in specifications.

Preliminary/Initial Design

Digital Demodulator (LGDT3303)

IF

MPEG Data

Antenna

Tuner (XC3028)

CVBS

Analog Video Decoder (TVP5150)

ITU656

Bridge (EM2883)

USB

SIF

CVBS

S-Video

Analog Audio Decoder (EMP202)

L/R Audio

Concept Study

Implementation Study

Concept Development

Implementation Development

Preliminary/Initial Design

P/I Design critique/analysis

Final Design

Knowledge and Relevant Coursework

This design makes use of the knowledge we have already acquired and that which will need to be acquired during the design process. The knowledge that the team has acquired in C++ will prove most useful in the execution of the task. Also, there has been research and knowledge acquired on behalf of Digiclips prior to the team's joining the project. The team will utilize the information that has already been found and applied by Digiclips on past endeavors and apply it

accordingly. The Team has some C++ experience but will have to gain more depth in experience and application of programming and databases in order to successfully complete the task. Open source information will also be a pivotal source of information so that the team can use common knowledge and research that has already been done and is available to the public. The application of the open source information will allow the team for focus more on tasks that require more detail and have not yet been solved as it pertains to the task.

Preliminary/Initial Design Critique & Analysis

This includes conducting multiple tests on the hardware and program through each of the 10 phases in order to ensure that the project is ready and capable to move on to the next phase of production and programming. Also to make sure it completely functions while adhering to the guidelines & restrictions that were made known to the team by Digiclips. This will also be the time period where adjustments will be made to increase the efficiency and user-friendliness of the RTV search engine. This time will also be used to add to the aesthetics of the program and make it more aesthetically pleasing.

Implementation Design and Analysis

A diagram of the 10 phases of the programming that were laid out by Digiclips, and these 10 phases are essentially a vague blueprint of how the team i s going to approach the solution as a whole.
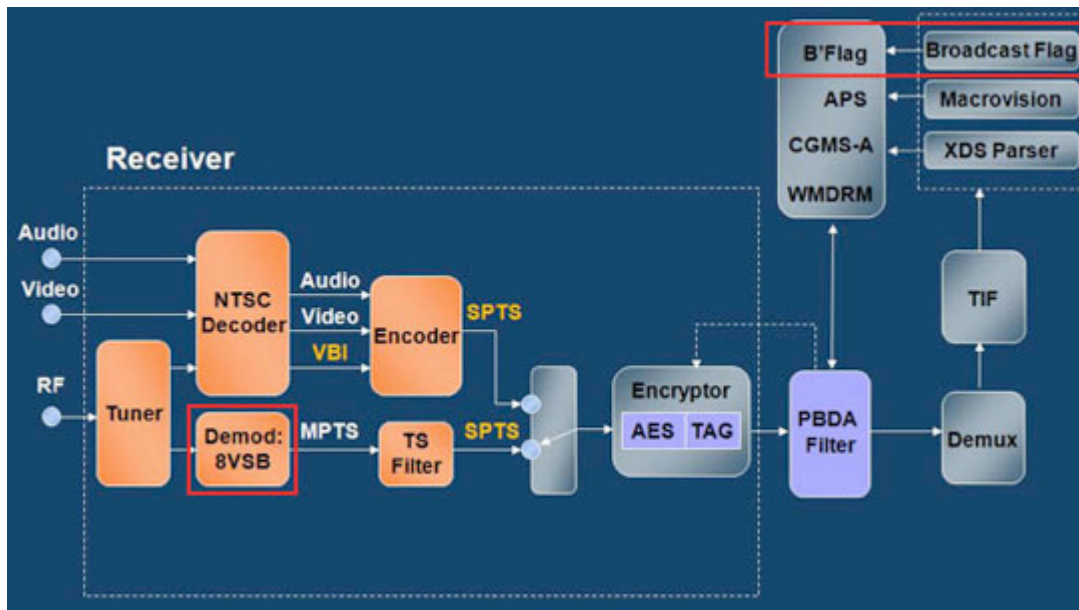
Communication

Here we will implement a communication system that will adhere to the rules that Digiclips has communicated with the team. The primary communication in the project will be communication between the hardware and software. Some of this communication will be attributed to packages and programming that come with the hardware and programming that has already been done by Digiclips in previous projects. The team will have to use programming knowledge in order to ensure that any other means of communication necessary are possible. The team will also have to use programming knowledge in order to make hardware and software communicate in accordance with the aforementioned 10 phases of programming.

Tasks and Deliverables

Team Member	Tasks
Jonathan Murphy	SQL Database Programming
Christopher Caesar	HVR board video integration
Bryant Beeler	HVR board audio Integration (Audio Mining)

Final Design & Milestones

Completion of the instantaneous radio television search engine is planned to be ready for approval by March of the following year.

Our milestones/deadlines are as follows:

Milestone	Deadline
Presentations	November 2011
Audio to Text Research Radio to Text Research SQL (Data Mining)	November 2011 – January 2012
Hardware and Software Acquisition	December 2011
Software Preliminary Design	December 2011
Integration Testing	January 2012
Software Final Design	February 2012

Final Design	March 2012

Cost and Resources

Materials	Prices
PCIe Slot (4GB RAM)	$399.00
Hauppauge HVR 2250 board	$60-$150
Miscellaneous	$30.00
Digital TV Cable (Antenna)	$ 10.00
Hauppauge HVR 2250 board (SDK)	Provided
C ++/.net/SQL Research materials Speech to text software	Open Source
Total	Approx $600

Conclusion

Digiclips is challenging the Howard University team to come up with a useful and practical solution approach. The group will be able to successfully complete the task and meet all of the requirements by following the 10 Phases of the programming that were laid out by Digiclips and are essentially a vague blueprint to the entire project.

The team will also utilize Free Open Source material on the net that can be customized in order to avoid "reinventing the wheel". The Benefit is that the team will be able to Expand programming and hardware knowledge while contributing to the company and solving a real world problem by presenting a practical solution. The timeline will be adhered by closely and the team will have completed the task in its entirety and be ready to present by March 31st 2012, ECE presentation day.

References:

1. www.digiclipsinc.com
2. http://www.hauppauge.com/site/products/data_hvr2250.html
3. http://forum.team-mediaportal.com/684620-post8.html
4. http://wiki.team-mediaportal.com/1_MEDIAPORTAL_1/14_Using_MediaPortal/3_TV/1_Watch_Live_TV/Fullscreen_TV/Actions_Menu/Subtitles