**ACKNOWLEDGEMENT**

We feel much obliged to the Head of Department of our Bachelors of Engineering in Computer and IT Program, Er Bibek Ropakheti, for providing the required initiatives and directives to exhibit our current level of knowledge in the real world application by means of assigning us this project. It is an honor to reflect our gratitude to our seniors who have helped paved the course for us in this project to make it more authentic.

We would also like to express our earnest appreciation to our amicable and yet inspiring Prof. Mrs. Divyeshwori Manandhar for supporting us with her genuine ideas and valuable suggestions. Among the list finally, our sincere unfeigned thanks to faculty members, staffs and friends that have helped us to scrutinize and prepare the current proposal.

Sincerely,  
Punita Kushwaha (BCE 14/316)  
Deepak Prasad Shah (BCE 14/328)  
Rohan Baidya (BCE 14/335)  
Deepa Bhandari (BCE 14/336)

**ABSTRACT**

Since the elementary levels, kids are taught about basic physical body parts, and among them lie our five most important sensory organs; Eyes, Nose, Ears, Tongue, and Skin – that lay a groundwork for human intelligence and communications. Albeit all of the mentioned organs are vital and significant for proper functioning of human body system, we are often disposed to make use of Vision and Sound more frequently to perceive, recognize and to use it as a means of communication.

As it is eminent that human mind is the nature’s state of art; moreover the concept behind running of whole body system is simple yet consists of astoundingly complex tissues and neural constructions. Furthermore all the sensory perceptions via receptor organs are transmitted in form of electrical signals to brain – where the brain; as a complex combination of logic circuit of nature; processes those input signals to generate a fruitful output by means of motor organs (i.e. Muscular movement, Speech, etc).

Current system of Visual Intelligence System (VIS) is intended to mimic the way we innately perceive things and communicate information. Keeping pace with the development of advanced hardware components for Audio and Visual interaction among user and the system; and as the computer now-a-days being far more capable of keeping afloat a gigantic cloud of information on top; using computer based recognition system has evolved to become more reliable and accurate.

**TABLE OF CONTENTS**

1. **ACKNOWLEDGEMENT------------------------------------------------------i**
2. **ABSTRACT----------------------------------------------------------------------ii**
3. **TABLE OF CONTENTS---------------------------------------------------- iii**
4. **LIST OF FIGURES------------------------------------------------------------iv**
5. **LIST OF TABLES--------------------------------------------------------------iv**
6. **ABBREVIATION--------------------------------------------------------------iv**
7. **INTRODUCTION-----------------------------------------------------------------1**
   1. **Literature Review-----------------------------------------------------------1**
   2. **Background ------------------------------------------------------------------2**
   3. **Open CV-----------------------------------------------------------------------3**
   4. **Sphinx--------------------------------------------------------------------------4**
   5. **Objectives---------------------------------------------------------------------5**
   6. **Applications-------------------------------------------------------------------6**
   7. **Entity Relationship Diagram----------------------------------------------7**
8. **PROBLEM STATEMENT-------------------------------------------------------8**
9. **METHODOLOGY-----------------------------------------------------------------9**
10. **EXPECTED OUTPUT-----------------------------------------------------------11**
11. **SCHEDULE------------------------------------------------------------------------12**
12. **CONCLUSION---------------------------------------------------------------------13**
13. **REFERENCE----------------------------------------------------------------------14**

**LIST OF FIGURES**

**LIST OF TABLES**

**ABBREVIATION**

* + 1. **API:** Application Programming Interface
    2. **CCTV:** Closed Circuit Television
    3. **CMU:** Carnegie Mellon University
    4. **viz:** namely
       1. **INTRODUCTION**
          1. **Literature Review**

Initiation of Visual Intelligence System takes from the time back in 1960s, when Woody Bledsoe, Helen Chan Wolf, and Charles Bison developed face recognition using computer. In 1996 A.D., Bledsoe 1968a was tested at Stanford Research Institute by Peter Hart in 1996 AD with 2000 photographs. Moreover, the result was majorly acceptable. In 1997, the system developed by Christoph von der Malsburg and graduate students of the University of Bochum in Germany, by United States Army Research Laboratory.

In 2001, Paul Viola and Michael Jones developed an efficient algorithm for Face Detection. Later, Object recognition is a process for identifying a specific object in a digital image or video. Object recognition algorithms rely on matching, learning, or pattern recognition algorithms using appearance-based or feature-based techniques. Common techniques include edges, gradients, Histogram of Oriented Gradients (HOG), Haar wavelets, and linear binary patterns. Object recognition is useful in applications such as video stabilization, automated vehicle parking systems, and cell counting in bioimaging.

* + - * 1. **Background**

Image Recognition has following models:

Extracted features and boosted learning algorithms

Bag-of-words models with features such as SURF and MSER

Gradient-based and derivative-based matching approaches

Viola-Jones algorithm

Template matching

Template matching is a technique for finding areas of an image that match (are similar) to a template image (patch).

Image segmentation and blob analysis

* + - * 1. **Open CV**
        2. **Sphinx**
        3. **Objectives**
        4. **Application**

**1. Biometrics:**

**2. Artificial Intelligence:**

**3. Autonomous Vehicle:**

**4. Species Identification:**

**5. Modeling objects:**

**6. Surveillance or Tracking:**

* + - 1. **DESIGN SCHEMATICS**
      2. **ENTITY RELATIONSHIP**
      3. **PROBLEM STATEMENT**

As of 2016, facial recognition is still not effective for most applications even though the accuracy has been substantially improved. Although systems are often advertised as having accuracy near 100%, this is misleading as the studies often uses much smaller sample sizes than would be necessary for large scale applications. Because facial recognition is not completely accurate, it creates a list of potential matches. A human operator must then look through these potential matches and studies show the operators pick the correct match out of the list only about half the time. This causes the issue of targeting the wrong suspect.

* + - 1. **METHODOLOGY:**

Since the features to be added in this system are unpredictable

* + - 1. **BUDGET STRUCTURE**

1. **WORKING SCHEDULE**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 7th Week  (Oct 10-16) | | | | | | | 6th Week  (Sep 12-18) | | | | | | 5th Week  (Sep 5-11) | | | | | | | 4th Week  (Aug 22-28) | | | | | | | 3rd Week  (Aug 15-21) | | | | | | | 2nd Week  (Aug 8-14) | | | | | | | 1st Week  (Aug 1-7) | | | | | | | Weeks  Work |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | Planning |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | Construction |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | Testing |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | Debugging |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | Integration |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | Finalizing |

Gantt Chart of Work Routine

1. **CONCLUSION**
2. **REFERENCE**

<https://phys.org/news/2016-12-facial-recognition-technology-real-world.html>

[site visited on 30th July 2017

<https://techcrunch.com/2016/11/16/facial-gesture-controls/>

[site visited on 30th of July 2017]]

<http://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/template_matching/template_matching.html> [site visited on 21st of July, 2017]

<http://www.audioanalytic.com/3-reasons-why-sound-recognition-is-the-next-big-thing-in-smart-home/> [site visited on 19th of July, 2017]

**Images:**

All the images extracted for the designs have been used under the GPL license.