

Image Driven Augmented Reality (IDAG) Using Image Processing

Ms. Vijayalakshmi Puliyadi
Assistant Professor
Computer Department
Fr.CRIT
Vashi, India

Deepak C. Patil
Computer Engineering
Fr.CRIT
Vashi, India

Pritam P. Pachpute
Computer Engineering
Fr.CRIT
Vashi, India

Haresh P. Kedar
Computer Engineering
Fr.CRIT
Vashi, India

Abstract - Real world object such as books, newspapers, posters etc. are not providing detailed realistic information about various objects, images and the content related to it as it may or may not be self-expressive. So, there is a need of technology which is next step to simply browsing the internet based on description, to intelligent application that can search or describe about what we see and provide realistic view of the environment. This concept can be fulfilled by using "Augmented reality". Our concept is much like what we see and our brain understands. We propose to implement it by capturing real world objects, images and augmenting the appropriate description using image searching.

Keywords – *Augmented Reality, Image Processing, Colour Blind People*

I INTRODUCTION

Data on the internet is growing at a drastic rate and due to this large information related to each and every object in the world, it can be easily available on the internet. But still finding the relevant information on the internet is very difficult as the search results are dependent on how the user describes the object in text format. Search engines provide various different search results even if one or two words in the query are changed or removed. Therefore, if user fails to describe the object correctly in manual text to search he may not get the correct result.

Moreover, we know that human has ability to easily memorize the objects that he sees in pictorial form and also to recollect it in the same form. So there is a need to provide relevant data directly to user related to the object without need of describing it in text.

Our idea of "IDAG" satisfies above mentioned need using camera-equipped, data enabled smartphone and "Augmented Reality". Smartphone screen can be used as the frame through which the user sees the environment and gets the information related to the object in that frame. We can use this concept for providing the information related to the image captured by smartphone in the form of multimedia data by replacing original image, which in turn gives realistic view as well as lead to easy understanding of the object. This concept is very helpful for students to understand the diagrams, architecture provided in the textbook with related multimedia displayed on screen. Also for any reader to get live feel of reading material such as newspaper, poster etc. by augmenting result on the user Smartphone screen.

II . METHOD

1. IMAGE DRIVEN SEARCH

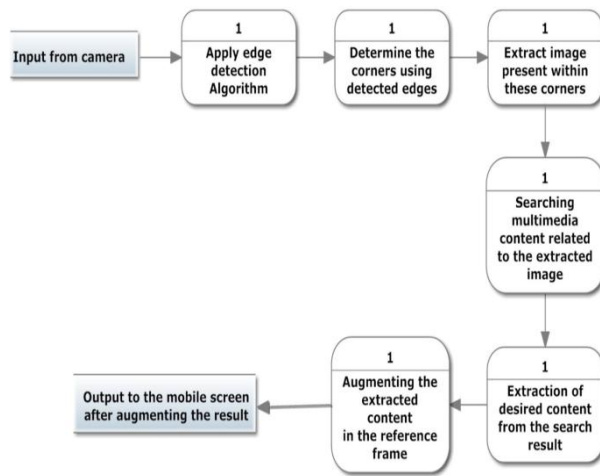


Figure 1 DFD level 1 for image driven search

A. EDGE DETECTION

Edge detection techniques are really important for object frame detection. The Canny Edge detection also known to many as the optimal detector can be used to detect the edges in an image.

Canny algorithm is preferred as it satisfies the three main criteria:

1. **Low error rate:** Meaning a good detection of only existent edges.
2. **Good localization :** The distance between edge pixels detected and real edge pixels have to be minimized.
3. **Minimal response:** Only one detector response per edge.

Thus we used the canny edge detection in our implementation.

Steps involved in the edge detection are :

1. Filter out any noise. The Gaussian filter is used for this purpose with kernel size of 3.

```
blur(src_gray, detected_edges, Size(3,3) );
```

2. Find the intensity gradient of the image.
 - 2.1 Apply a pair of convolution masks (in x and y directions

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix}$$

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix}$$

- 2.2 Find the gradient strength and direction with

$$G = \sqrt{G_x^2 + G_y^2}$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right)$$

3. Non-maximum suppression is applied. This removes pixels that are not considered to be part of an edge. Hence, only thin lines (candidate edges) will remain.
4. Hysteresis: The final step. Canny does use two thresholds (upper and lower).

4.1 If a pixel gradient is higher than the upper threshold, the pixel is accepted as an edge.

4.2 If a pixel gradient value is below the lower threshold, then it is rejected.

4.3 If the pixel gradient is between the two thresholds, then it will be accepted only if it is connected to a pixel that is above the upper threshold. Canny recommended a upper:lower ratio between 2:1 and 3:1. [6]



Figure 2. Scanned Image through smartphone

We used the OpenCV for image processing in IDAG. OpenCV method to perform canny edge detection is **Imgproc.canny(gray,intermat,150,160);**

Where,

1. Gray : Source image, grayscale
2. Intermat: Output of the detector (can be the same as the input)
3. lowThreshold: The low threshold value.
4. highThreshold: the high threshold value.



Figure 3. Image after Canny edge detection.

B. DETERMINING THE CORNERS

To extract the image of an object it is very much necessary to find the corners like left top and right bottom corners which are sufficient to define image coordinates. The corners can be located by finding the intersection point of horizontal and vertical edges.

In case of non-rectangular objects, the bounding boxes around the largest possible object can be drawn. The corners of that box is now considered for extraction.

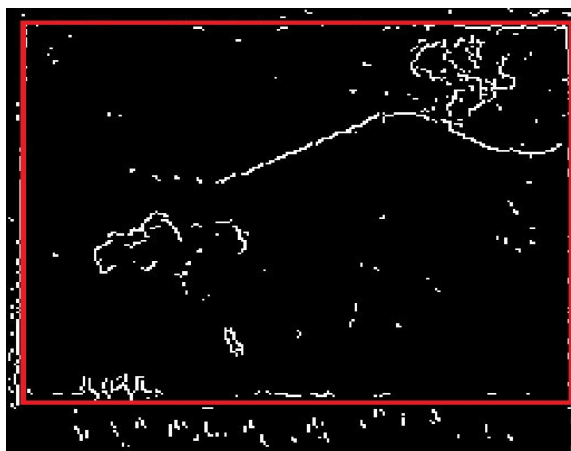


Figure 4. Detection of rectangular frame of object with the help of corner detection

C. EXTRACTION OF IMAGE

This phase includes the image extraction within corners which we have calculated in phase B. Various libraries like OpenCv which are available for Androidbased Smart phones can be used for image extraction.



Figure 5. Extracted image.

D. RELEVANT MULTIMEDIA CONTENT SEARCH AND EXTRACTION

We used imgur.com for image hosting Application Program Interface (API) and Bing search API to retrieve the related information which include steps as below [7],

Step 1: Upload image on imgur.com

Step 2: The direct link of uploaded image is extracted and it is given as input to Bing Search API

Step 3: Execute a search:

Execution of the query related to the image search includes method calls such as

```
.image("direct link", null, null, null, null, null);
```

initiates a new search, where query supplies the search term.

Step 4: Result Extraction:

With the search we can extract various results like Video, Image, Text etc to extract the user favourable multimedia content.

E. AUGMENTING THE RESULT

Step 1: Camera Pose Estimation and Scene Registration

Once new features such as edge shape of the object are detected from the camera image, their corresponding 3D object's positions are computed from the projective mapping given by the initial geometric settings. The camera pose can then be successively estimated from the 2D-3D projective relationship of the feature points

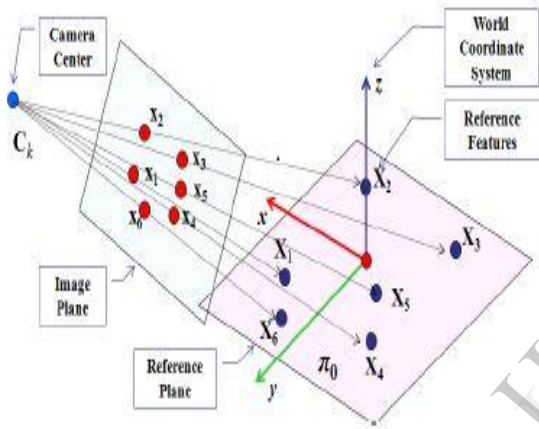


Figure 6. Projective mapping between the reference plane and its image

Step 2: Feature Tracking and New Feature Detection

The detected features are successively tracked by Tracker every frame, and the camerapose is computed based on the 2D-3D projective mapping. Therefore, when the number of features in the detected feature list drops below a predefined threshold, we detect new features and their corresponding 3D coordinates are recomputed based on the camera projective mapping.

We assume that all detected features are located on the reference plane, so that their 3D coordinates can be computed from a plane-to-plane projective homography which greatly reduces computational cost of feature depth estimation required for acquisition of 3D coordinates of new features [3].

2. APPLICATION FOR COLOUR BLIND PEOPLE

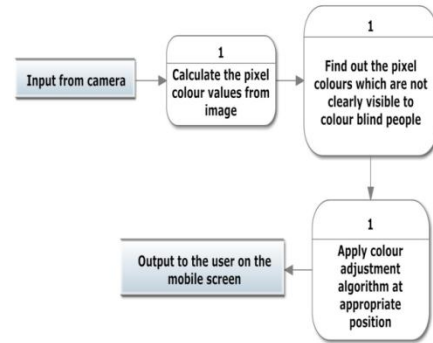


Figure 7 DFD Level 1 Application for colour blind people

“Augmented Reality” (AR) which is base concept to implement the project IDAG, has various application like one which provides environment to colour blind people by adjusting RGB values (pixel colour values) so that they can also see the surrounding as normal human being. [4] Fig 5 gives the DFD for implementing our proposed concept to application for colour blind people. This feature can be implemented as:

A. CALCULATING PIXEL VALUES

The image is taken as input. The pixel values of an input image are calculated in RGB format

B. FINDING PIXELS WHICH ARE NOT VISIBLE TO USER

According to type of colour blindness of user, finding required pixel to be fixed in an image

C. ADJUSTING REQUIRED PIXEL VALUES

The founded pixel are applied to colour adjustments algorithm and the pixel values are adjusted such that colour blind user can differentiate with all colours. This image is given as an output to the colour blind use

III. CONCLUSION

In this paper we proposed a concept of searching real world object related multimedia data using image of it and augmenting the searched result on camera equipped and data enabled Android smartphone screen. Also the application for colour blind people such that they can see the resulted media content in distinguishable colours by adjusting RGB pixels values. This concept can be applied to many applications like e-learning, live newspaper.

IV. REFERENCES

- [1] ***Video-Based In Situ Tagging on Mobile Phones*** by Wonwoo Lee, Youngmin Park, Vincent Lepetit, and Woontack Woo. Circuits and Systems for Video Technology, IEEE Transactions on Volume:21, Digital Object Identifier: 10.1109/TCSVT.2011.2162767
Publication Year: 2011, Page(s): 1487 – 1496
- [2] ***Developing Mobile Mixed Reality Application Based on User Needs and Expectations*** by Amandeep Dhir, Thomas Olsson, Said Elnaffar. Innovations in Information Technology (IIT), 2012 International Conference on Digital Object Identifier: 10.1109/INNOVATIONS.2012.6207780
Publication Year: 2012, Page(s): 83 – 88
- [3] ***Real-time Camera Pose Estimation Based on Planar Object Tracking for Augmented Reality Environment*** by Ahr-Hyun Lee, Seok-Han Lee, Member, IEEE, Jae-Young Lee, and Jong-Soo Choi, Member, IEEE. Consumer Electronics (ICCE), 2012 IEEE International Conference on Digital Object Identifier: 10.1109/ICCE.2012.6162000
Publication Year: 2012, Page(s): 516 – 517
- [4] ***Traffic Indication Symbols Recognition with Shape Context*** by Kai Li, Weiyao Lan. Computer Science & Education (ICCSE), 2011 6th International Conference on Digital Object Identifier: 10.1109/ICCSE.2011.6028771
Publication Year: 2011, Page(s): 852 – 855
- [5] ***Using camera equipped mobile phones for interacting with real world objects*** by Michael Rohs, Beat Gfeller
- [6] http://docs.opencv.org/trunk/doc/tutorials/introduction/android_binary_package/dev_with_OCV_on_Android.html
- [7] <http://www.bing.com/developers/s/APIBasics.html>