MOVING OBJECT RECOGNITION USING BACKGROUND SUBTRACTION AND IMAGE PROCESSING TECHNIQUES

[1]MERVIN K, [2]SARALA B

[1] Master of Engineering in Applied Electronics, Sri Venkateswara College of Engineering, Sriperumbudur
[2] Assistant Professor, Department of Electronics and Communication Engineering , Sri Venkateswara College of Engineering, Sriperumbudur

Abstract: A visual prosthesis that applies electrical stimulation to different parts of the visual pathway has been proposed as a viable approach to restore functional vision. However, the created percept is currently limited due to the low-resolution images elicited from a limited number of stimulating electrodes. Thus, methods to optimize the visual precepts providing useful visual information are being considered. Image-processing strategies based on a novel background subtraction technique to optimize the content of dynamic scenes of daily life are given to the prosthetics. Psychophysical results showed that background reduction, or background reduction with foreground enhancement, increased response accuracy compared with methods that directly merged pixels to lower resolution. By adding more gray scale information, a background reduction/foreground enhancement strategy resulted in the best performance and highest recognition accuracy. Further development of image-processing modules for a visual prosthesis based on these results will assist implant recipients to avoid dangerous situations and attain independent mobility in daily life.

Keywords: Visual Prosthetics, functional vision, Background reduction, foreground enhancement.

I. INTRODUCTION

In the normal visual pathway, light travels through the tear film, cornea, aqueous, pupil, lens, and vitreous, to activate the light sensitive photoreceptors and set up the trans-synaptic connections of the retina. The brain does all of the complex image processing, while the eye functions as the biological equivalent of a camera.

Millions of people worldwide lose their photoreceptors either due to retinal degenerations (e.g. retinitis pigmentosa (RP)) or age-related macular degeneration (AMD). The feasibility of an implantable retinal prosthesis that would partially restore vision by direct electrical stimulation of retinal neurons is supported by several studies. Morphometric analyses in post-mortem eyes with almost complete photoreceptor loss either due to RP or AMD has shown as many as 90% of the inner retinal neurons remain histological intact. When there is a gross destruction of the eye like phthisis bulbi, restoration of vision is not possible. But in conditions where blindness is due to a disease in the photoreceptors in the retina like Retinitis Pigmentosa (RP) as shown in figure 1 or Age Related Macular Degeneration (AMD) as shown in figure 2 wherein the neural connections are intact, in cortical tumours and lesions involving the visual pathways or occipital cortex wherein the photoreceptors are intact, there exists a scientific possibility wherein a device can be implanted at any location in this pathway to set up a neuronal electrical stimulation. Such a device is known as a visual prosthesis or bionic eye and the vision created is known as artificial vision. Incidentally, RP is the leading inherited cause of blindness, with 1.5 million people affected worldwide. AMD seen in adults over 65 years is the leading cause of visual loss with 700,000 new cases diagnosed annually in the USA and 10 % of whom become legally blind each year. Moreover, these are a group of blind people who have had good vision all their life and have become blind at a time when all faculties are on the decline and daily survival requires vision.

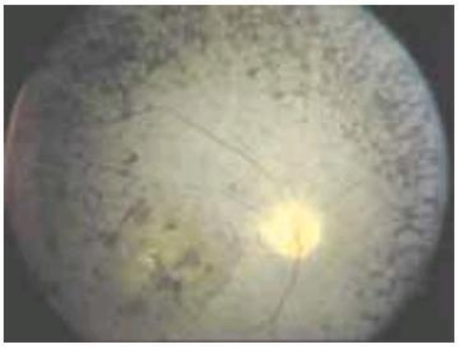


Figure 1: Retinitis Pigmentosa

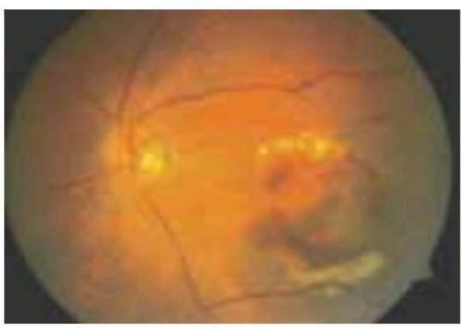


Figure 2: Age related Macular Degeneration

With increasing research advances and clinical trials of visual prostheses, there is significant demand to better understand the perceptual and psychophysical aspects of prosthetic vision. In prosthetic vision a visual scene is composed of relatively large, isolated, spots of light so-called ‘‘phosphenes”, very much like a magnified pictorial print. The utility of prosthetic vision has been studied by investigators in the form of virtual–reality visual models (simulations) of prosthetic vision administered to normally sighted subjects. Prosthetic vision is built upon phosphenes. A phosphene is any visual sensation caused by means other than stimulation of the visual system by light. This encompasses phosphenes elicited by mechanical forces and magnetic stimulation but in the present context, the focus is on phosphenes elicited by electric stimulation. Particularly, a phosphene will be referred to as a single, elementary spot of light in the visual field unless explicitly stated otherwise. This means that a cluster of dots elicited will be addressed as a cluster of phosphenes, and a merged patch of light from multiple electrodes will be addressed as a combinatorial effect of multiple phosphenes as shown in figure 3.

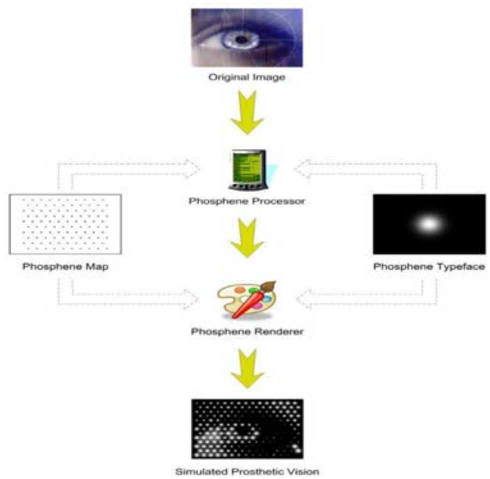


Figure 3: Phosphene Model

Taking the idea of phosphenes as the elementary visual components, many investigators have implemented simulations (or visual models) of the anticipated form of restored vision (SPV) by extrapolating from the description of singular phosphenes in the literature to a visual field composed of a large collection of phosphenes; this is the expected form of prosthetic vision to be provided in future visual prostheses.

II. BACKGROUND SUBTRACTION TECHNIQUE

There are no effective clinical treatments to restore vision for some retinal diseases such as age-related macular degeneration and retinitis pigmentosa. Implanting a visual prosthesis has been proposed as a viable approach to restore partial vision to blind patients suffering from these diseases. The perception of spots of light, called phosphenes, are elicited by electrically stimulating different parts of the visual pathway (retina, optic nerve, or cortex). Over recent decades, several research groups have developed different types of visual prosthetic devices and successfully implanted them in blind patients. Although visual prostheses have gained significant development and continue to achieve encouraging improvement, some engineering challenges, such as electrode fabrication, power consumption, and long-term viability, remain to be overcome before microelectronic high density electrode implants can be realized. Consequently, visual perception generated by a limited number of stimulation contacts is still poor relative to normal vision. Methods to optimize the image quality presented by such a limited number of phosphene dots to maximize visual percepts are currently being considered.

Background subtraction is process of extracting foreground objects from maintained background model. A foreground object is any entity that detected by producing difference of the every frame of sequence to background model. This result can be further used for tracking targets, motion detection. Background subtraction further divides into parametric and non-parametric background subtraction. There are different background subtractions techniques have been proposed in literature. The background model can be static or dynamic. The flowchart for Background subtraction is shown in figure 4. Dynamic background model is one in which the background of scene may contain moving objects in outdoor environment, Pixel-based and block based are two major kind of approached are for background Subtraction. To construct a statistical representation of background scene non-parametric statistical Modelling of pixel process is used. The different challenges that have to face to construct a good background subtraction algorithm are robustness against the changes in illumination and shadow detection.

Background subtraction is a widely used approach for detecting moving objects in videos from static cameras. The rationale in the approach is that of detecting the moving objects from the difference between the current frame and a reference frame, often called the "background image", or "background model". As a baric, the background image must be a representation of the scene with no moving objects and must be kept regularly updated so as to adapt to the varying luminaries conditions and geometry settings. More complex models have extended the concept of "background subtraction" beyond its literal meaning. Several methods for performing background subtraction have been proposed in the recent literary. All of these methods try to effectively estimate the background model from the temporal sequence of the frames. However, there is a wide variety of techniques and both the expert and the newcomer to this area can be confused about the benefits and limitations of each method. This paper provides a thorough review of the main methods (with inevitable exclusions due to space restrictions) and an original categorisation based on speed, memory requirements and accuracy.

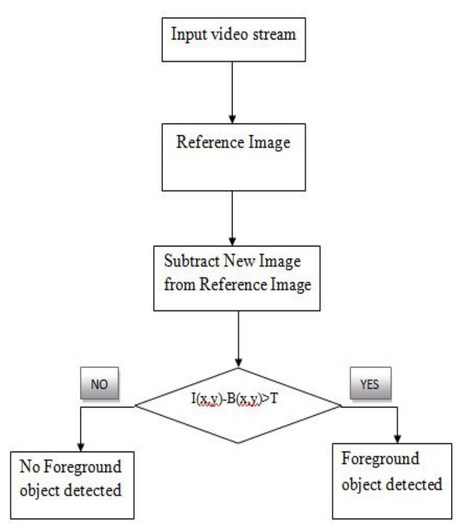


Figure 4: Flow chart of back ground subtraction

Background subtraction is one of the key techniques for automatic video analysis, especially in the domain of video surveillance. Although its importance, evaluations of recent background subtraction methods with respect to the challenges of video surveillance suffer from various shortcomings. To address this issue, we first identify the main challenges of background subtraction in the field of video surveillance. We then compare the performance of nine background subtraction methods with post-processing according to their ability to meet those challenges. Therefore, we introduce a new evaluation data set with accurate ground truth annotations and shadow masks. This enables us to provide precise in depth evaluation of the strengths and drawbacks of background subtraction methods. A common method for search-space reduction and focus of attention modelling in video analysis is background subtraction (BS). The term BS covers a set of methods that aim to distinguish between foreground and background areas in video sequences utilizing a background model. Over the past years, various BS methods have been developed, with each of them having its own characteristic, strength and weakness. Evaluation allows identifying those characteristics and helps to focus on the remaining problems. One reason might be the huge effort involved in generating qualitatively high ground truth (GT) data of natural video sequences. Hence, some evaluations only use a few labelled frames or judge the performance at object level, which is considerably easier. However, evaluation at pixel-level provides more insight into strengths and weaknesses. There exist several techniques to overcome manual GT annotation.

III. VISUAL BACKGROUND EXTRACTOR-GAUSSIAN MIXTURE MODEL

RGB colour video images were resized to a 480X480 resolution, converted to gray scale, and then adjusted for uniform illumination. In general, the image processing stage of the visual prosthesis adjusts the image resolution by combining a set number of pixels into a single output pixel for stimulating the tissue interface array and is called Merging Pixels to Low Resolution (MPLR). The small electrode number leads to a huge loss of information when presenting a real-life scene. Increasing the contrast between a moving object(s) (i.e. the foreground) and the relatively static or slow moving parts of the scene (i.e. the background) can enhance the perception of the main information. Therefore, moving objects in a dynamic scene need to be automatically detected and precisely separated from the surrounding information, i.e. a series of foreground segmentation images must be generated first from the original video. A common motion detection technique for distinguishing the foreground from the background in computer

International Journal of Scientific Engineering and Applied Science (IJSEAS) - Volume-1, Issue-8,November 2015
ISSN: 2395-3470
www.ijseas.com

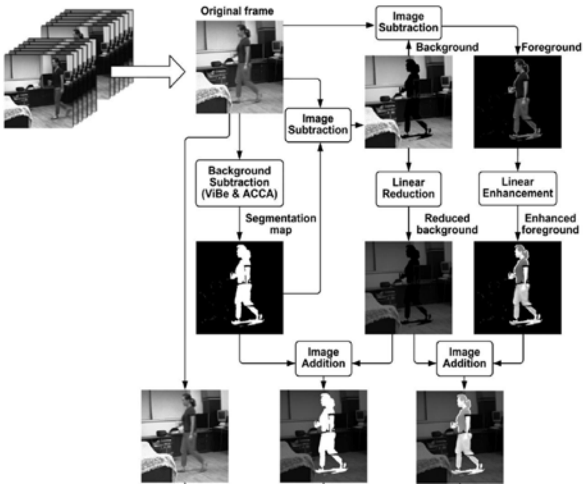


Figure 5: Image processing strategy for BS

vision is background subtraction (BS). We applied a novel BS algorithm for foreground extraction and developed a post-processing method for optimizing the segmentation. Two image-processing strategies based on BS segmentation were used to increase the contrast between the moving foreground and background as shown in figure 5, and then compared with direct MPLR processing.

VISUAL BACKGROUD EXTRACTOR

BS regards the foreground as the difference between the current image and a reference background image, often called the ‘‘background model’’. Several techniques have been used, such as the single Gaussian model or temporal median filtering that are simple models but extremely sensitive to changes in dynamic scenes resulting from lighting and/or extraneous events. The Gaussians Mixture Model and its enhanced version in real-time tracking provide good model accuracy, albeit at the price of increased algorithm complexity and computational time. A recent proposed universal BS technique called visual background extractor (ViBe) outperforms other techniques, not only in terms of computational speed and classification accuracy, but also the ability to solve problems arising from the nature of video surveillance. Due to its excellent performance and almost parameterless procedure, we used ViBe for foreground segmentation. ViBe compares current input pixels with a background model by randomly selecting neighbouring pixels.

The ViBe background model is initialized with the neighbourhood of each observed pixel from the first frame based on the assumption that a pixel and its neighbours share a similar distribution in the spatial domain. Suppose at the starting time in a video sequence that $BG(x,y), N_G(x,y) and\ (x,y)_m$ respectively represent the initial background model, the spatial neighbourhood of the pixel, and pixel samples in the model from the first frame. The initial background model is defined as:

$$BG(x,y) = \{(x,y)_m \mid (x,y)_m \in N_G(x,y), m = 1,2, \dots\dots N\} \quad (1)$$

$(x,y)_m$ is randomly selected in $N_G(x,y)$ according to uniform probability, and N is the total number of samples. When $t = k\ (k \neq 1)$, each $(x,y)^k$ in frame $f(x,y)^k$ is estimated via the last established background model $BG^{k-1}(x,y)$ at this location. In order to judge whether the pixel $(x,y)^k$ belongs to the background, the following equations 2 and 3 are used:

$$S_R^k = \{(x,y)_R^{k-1} \mid \|(x,y)^k - (x,y)_R^{k-1}\| \leq R, (x,y)_R^{k-1} \in BG \quad (2)$$

$$\#\{S_R^k\} \leq \#_{min} \quad (3)$$

where R is a spherical searching scope of $(x,y)^k$. S_R^k is the pixel set in which the distance between each pixel $(x,y)_R^{k-1}$ in $BG^{k-1}(x,y)$ and (x, y)k is less than R. If the cardinality of S_R^k, i.e. $\#\{S_R^k\}$, is less than a threshold $\#_{min}$, $(x,y)^k$ is classified as the background pixel. Otherwise, the pixel belongs to foreground. Classification processing ends when $\#_{min}$ matches are found. Due to the spatiotemporal property of background subtraction, a memoryless update policy was adopted for discarding some pixel samples in the model. The probability P of preserving a sample present in the model at time t after the update of the pixel model is given by (N - 1)/N, which is denoted as equation 4

$$P(t,t+dt) = \frac{N-1^{(t+dt-t)}}{N} \qquad (4)$$

Thus, the life span of a pixel in the background model is time-independent and decays exponentially. In addition, a random time sub-sampling factor, which selects new samples from an established background $f_B^k(x,y)$, is used for updating the background model over time. The random sampling mechanism propagates background pixel samples to guarantee spatial consistency and to update the background pixel model hidden by the foreground.

IV SIMULATION RESULT

MATLAB is a high-level language and interactive environment for numerical computation, visualization, and programming. Using MATLAB, you can analyze data, develop algorithms, and create models and applications. The language, tools, and built-in math functions enable us to explore multiple approaches and reach a solution faster than with spreadsheets or traditional programming languages, such as C/C++ or Java.

Various samples are taken and the Background subtraction for static or stationary images is performed. **Background subtraction**, also known as Foreground Detection, is a technique in the fields of image processing and computer vision wherein an image's foreground is extracted for further processing (object recognition etc.). Generally an image's regions of interest are objects (humans, cars, text etc.) in its foreground. Background subtraction is a widely used approach for detecting moving objects in videos from static cameras. The rationale in the approach is that of detecting the moving objects from the difference between the current frame and a reference frame, often called "background image", or "background model". This technique is used for the post processing method for the background subtraction for videos.

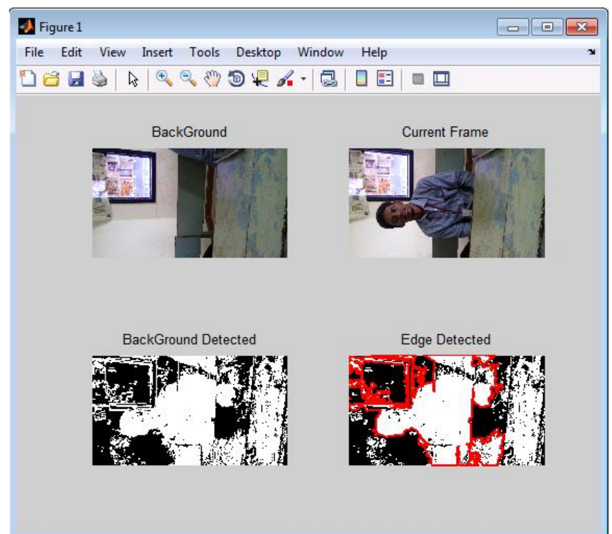
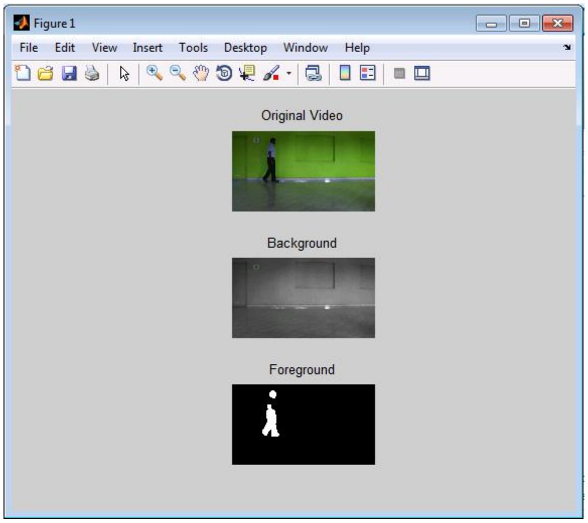


Figure 6: Background subtraction for images

In figure 6, the original image is taken and the background reference image is also taken. Using these images the background of the image is detected along with the edge details. Background subtraction can also be performed for videos. This technique uses visual background extractor technique using the Gaussian model techniques. Figure 7 and 8 shows the background subtraction or fore ground enhancement for the video at different positions.

Frames are generated from the video and the frames can be stored in the disk so that they can be programmed in the prosthetics designed. The dialog box for obtaining individual frames is shown in figure 9. Figure 10 and figure 11 shows the different visuals of the video which are converted into frames



and the mean graylevel showing the RGB components are also displayed.

Figure 7: Background subtraction for video at the initial position

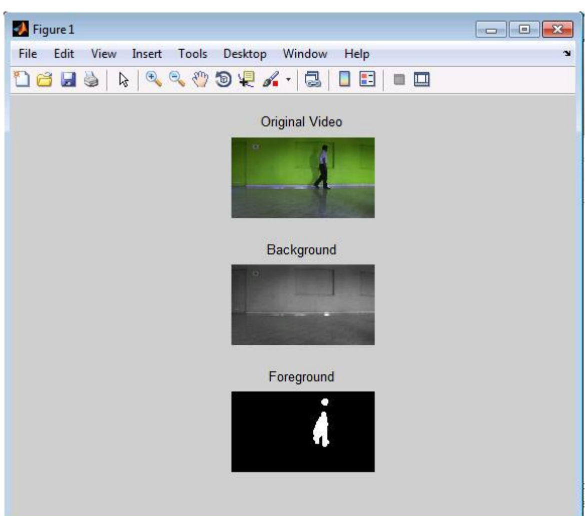


Figure 8: Background subtraction for video at a different position

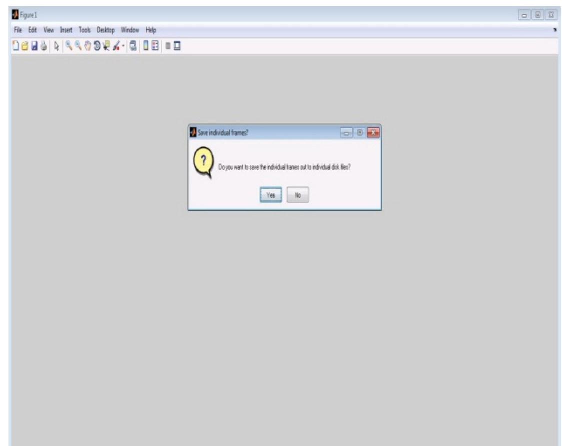


Figure 9: Video analysis by frame generation

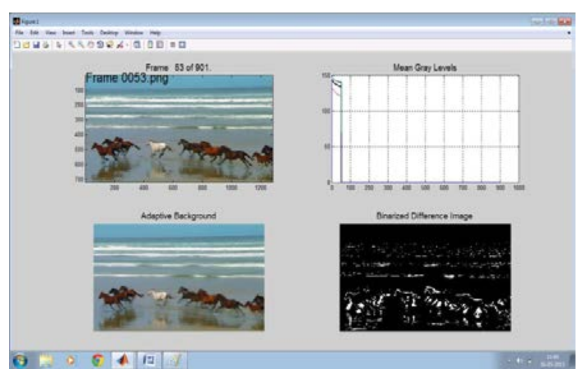


Figure 10: Frames generated for the video and the mean graylevel (frame 53)

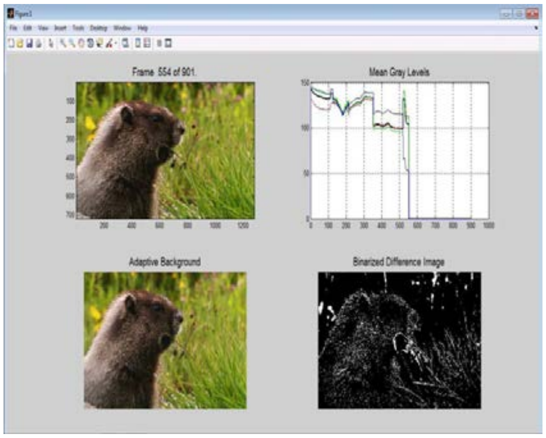


Figure 11: Frames generated for the video and the mean graylevel (frame 554)

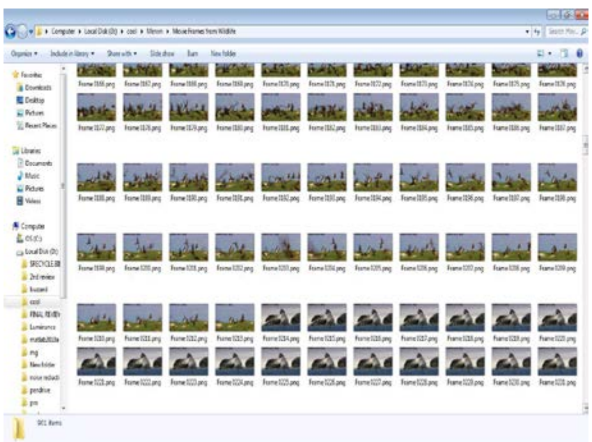


Figure 12: Frames of the video stored in the destination folder

After the completion of extracting the frames from the original video the frames are stored in the specified destination folder as shown in figure 12.

V.CONCLUSION

Visual prosthetic research has made significant progress since the first human tests over half a century ago, and its focus has already shifted from one of feasibility to one of optimizing the visual presentation and technical development. Exploring effective image processing strategies to optimize information content and to improve the perception of moving objects is focussed here. This project concentrates on Visual Background extraction using Gaussian modelling and the outputs are obtained using MATLAB software.

For future work, it is hoped that background-subtraction based image processing strategies will encourage further development of image-processing modules for a visual prosthesis that will assist implant recipients to avoid dangerous situations and attain independent mobility in daily life.

REFERENCES

1. Angelica Perez Fornos, Alexandre Pittard and Avinoam B.Safran (2008), '*Visual information required to achieve simple pointing and manipulation tasks*', in Vision Research 48, pp.1705–1718.

2. Bharti & Tejinder Thind (2013), '*Background Subtraction Techniques-Review*', in International Journal of Innovative Technology and Exploring Engineering (IJITEE), Vol. 2, No. 13, pp.29–32.

3. Bin Wang, Yu Liu, Wei Xu, Wei Wang and Maojun Zhang (2014), '*Background subtraction using spatiotemporal condition information*', in Optik 125, pp.1406– 1411.

4. Jing Wang, Yanyu Lu, Liujun Gu, Chuanqing Zhou & Xinyu Chai (2014), "*Moving object recognition under simulated prosthetic vision using background-subtraction-based image processing strategies*," in Information science 227, pp.512–524.

5. Mark S. Humayun and James D. Weiland (2003), '*Visual perception in a blind subject*

with a chronic microelectronic retinal prosthesis', in Vision Research 43, pp.2573–2581.

6. Massimo Piccardi (2004), '*Background subtraction Technique: a review'*, in IEEE International Conference on Systems, Man and Cybernetics, pp.3099-3104.

7. Spencer C. Chen, Gregg J. Suaning, John W. Morley and Nigel H. Lovell (2001), '*Visual models of phosphenes'*, in Vision Research 49, pp.1493– 1506.

8. Wenbing Tao, Yicong Zhou, Liman Liu, Kunqian Li, Kun Sun and Zhiguo Zhang (2014), '*Spatial adjacent bag of features with multiple super pixels for object segmentation and classification'*, in Information science 227, pp.373– 385.