EYESFIRST GAME – Eye Tracking API

Abstract:

Eyes first game is an Eye Tracking API game published by Microsoft. this paper proposes the working of the Eyes first game with eye tracking compatible devices by Digital Image Processing algorithms. Digital image processing represents a relatively new but very dynamic domain. Its importance is still growing due to the rapid development in related computer equipment. In this paper basics concepts are introduced, and a survey of problems and methods is presented. Starting with techniques used for the capture of data, it continues with techniques applied to coding new data, image enhancement and pattern recognition. Special attention is given to pattern recognition techniques which provide the basics for classification of data according to established criteria.

Introduction:

Eye tracking technology is the process of measuring either the point of gaze (where one is looking) or the motion of an eye relative to the head. An eye tracker is a device for measuring eye positions and eye movement. Eye trackers are used in research on the visual system, in psychology, in psycholinguistics, marketing, as an input device for human-computer interaction, and in product design. Eye trackers are also being increasingly used for rehabilitative and assistive applications (related for instance to control of wheelchairs, robotic arms and prostheses). There are number of methods for measuring eye movement. The most popular variant uses video images from which the eye position is extracted.

Eye tracking:

They say the eyes are the window to the soul, but they're also the gateway to knowledge about how people gather information and what influences their actions and decisions. The concept of studying where we look has been around since the 1800s, and today modern eye tracking technology allows us to gather an unrivaled amount of insight into the human mind.

Technically speaking, eye tracking is the process of measuring eye movements to determine where a person is looking, what they are looking at, and for how long their gaze is in a particular spot. Because our eyes are one of the primary tools, we use for decision making and learning eye tracking is commonly used by researchers and businesses looking to study human behavior because it's the only way to accurately and objectively measure and understand visual attention.

How does an eye tracker work?

An eye tracker uses invisible near-infrared light and high definition cameras to project light onto the eye and record the direction it's reflected off the cornea. Advanced algorithms are then used to calculate the position of the eye and determine exactly where it is focused. This makes it possible to measure and study visual behavior and fine eye movements, as the position of the eye can be mapped multiple times a second. How quickly an eye tracker is able to capture these images is known as its frequency. A recording can also be made of the scene a person is looking

at and using eye tracking software it's possible to produce a visual map of how the person viewed elements of the scene.

In this paper we will explore the range of eye trackers available, what types of research each is most suited to, how to understand and interpret the results of your eye tracking study, and the unique benefits of using eye tracking in research.

What types of eye trackers are there?

While the general principal of how they work is the same, there are several different eye tracking device types available and the one most appropriate for the user depends on the nature of their research. The main groups are:

Screen based – These are stand-alone, remote devices which either come as an individual unit or a smaller panel which can be attached to a laptop or monitor.

Wearable – These include eye tracking glasses and virtual reality (VR) headsets with integrated eye tracking.

Webcam – Webcam eye trackers don't have sensors or specialized cameras, they are solely comprised of the webcam device attached or built-in to a computer.

Screen based eye trackers are mostly used for research where the participant interacts or is exposed to the stimuli on a screen. These remote eye trackers offer broad sampling rates, and those with a high frequency can provide a large amount of data and a very high level of detail relating to the movement of the eye. Screen based eye trackers with a large tolerance for head movement are commonly used when studying people with certain medical conditions and infants who are unable to control their movements.

Eye tracking glasses are ideal for studying behavior in real-world situations such as browsing the aisles of a supermarket, playing sport, navigating the subway system, consuming media in the home, human interaction, or working in a factory etc. They allow natural movement and are unobtrusive, which means the wearer can do what they normally would in the situation. They also contain a built-in scene camera and microphone to record the environment.

VR headsets with integrated eye tracking allow situational interactions to be examined in multiple environments without the need for the environment to be physically present. This is useful for skills and safety training in places that would otherwise be too risky to be in, or are too difficult or expensive to set up. This method is also used to test things like building or shop design without the expense and time needed to physically create the place. This can be a very cost-effective way to test many scenarios across multiple participants in an efficient way.

Webcam eye tracking uses the built-in or external webcam attached to a laptop or monitor to collect information on where the person is looking. This method does not use infrared light beams or specialized cameras, instead it relies on the image generated from the webcam. An algorithm is then used to calculate the position of the head and eyes, and from that the direction of the eyes is correlated to an image on the screen. While the depth and accuracy of information you can get using this method is somewhat limited, webcam eye tracking enables large-scale

studies and a fast turnaround which is ideal for quantitative research. This method is commonly used early in a design process, such as A/B testing of website or product design.

What are the benefits of eye tracking in research?

One of the main benefits of eye tracking is that it's the only method which can be used to objectively and accurately record and analyze visual behavior. It would be impossible to ask someone scanning the aisles of a supermarket to recall, let alone quantify, the amount of time they spent looking at every item, or even exactly where they looked or what advertisements they noticed most.

Eye tracking allows researchers to study the movements of a participant's eyes during a range of activities. This gives insight into the cognitive processes underlying a wide variety of human behavior and can reveal things such as learning patters and social interaction methods. It also allows for the screening of atypical neurodevelopment and cognitive or perceptual disabilities. Eye tracking technology also provides a way for young children and those with neurological impairments, who are unable to explain their thought processes, to participate in studies.

Eye tracking:

- **Reveals subconscious behavior** – researchers can get insight into behaviors we carry out instinctively.
- **Provides unbiased, objective, and quantifiable data -** it removes the need to try and remember or explain where you looked and prevents study participants assuming details and giving incorrect information.
- **Allows for natural behavior –** eye trackers are unobtrusive and allow tasks to be carried out as normal.
- **Is versatile and mobile –** it can be used in almost any environment and setting.
- **Provides a high level of detail –** depending on the device and software, the results can offer a very high level of granularity for deep analysis.
- **Offers real-time information –** with live streaming you can see the person's gaze immediately.
- **Is explanatory –** it can depict processes and actions that are hard to articulate or explain.
- **Offers a visual representation -** AOI's, heat maps' and 'gaze plots' show the eye tracking results and how people have interacted with an environment or responded to a stimuli.
- **Adds value to other biometric data -** it can be combined with EEG, ECG, EMG, GSR NIRS and more. Eye tracking can enhance the use of these devices by providing additional information about what led to the physiological responses.

What are the main applications of eye tracking in research?

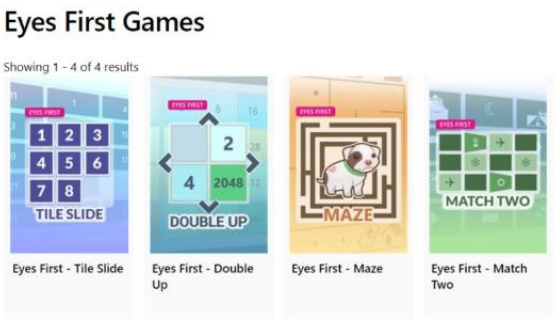
While there are almost endless fields in which this technology and research methodology could be applied, there are several which use eyes tracking extensively. They include:

- **Market Research –** Eye tracking offers detailed and unbiased information about consumer behaviors and decision-making processes. It allows market researchers and

brand owners to study the process consumers undertake when viewing and selecting a product. It shows them what elements naturally attracted the most attention and what areas were ignored. Unlike surveys or questionnaires, eye tracking details authentic behavior which is useful when designing advertising, branding, packaging and product placement.

- **User Experience -** There is no better way to test user experience than to view it through the eyes of the user. Eye tracking can be used to study the way in which platforms and services are used and how effectively they deliver their goals. Eye tracking can reveal design flaws and even methods of use which may not have been evident in the creative process.
- **Academic Research –** By studying visual behavior it's possible to get valuable insight into development, learning patterns and signs of cognitive dysfunction or disease such as those seen in patients with Alzheimer's, Parkinson's, schizophrenia, autism, depression, brain injury, and many more. Similarly, dyslexia and other reading or learning difficulties can be identified and studied by using eye tracking.
- **Professional Performance –** Eye tracking can give valuable insight into the methods in which tasks are performed and processes implemented. Eye tracking is used by businesses to identify safety risks, operational inefficiencies, and to streamline training; all saving substantial amounts of time and improving productivity. Similarly, sports professionals and coaches can use eye tracking to identify skills and strategies to improve performance. Eye tracking gives a unique snapshot into tasks and actions which are carried out quickly and often subconsciously. With the video playback and gaze mapping, it's possible to break complex and rapid processes into digestible information which can be studied and turned into training material.

The Eyes First games—Tile Slide, Match Two, Double Up, and Maze—aren't much in terms of actual games, as they're just twists on existing apps. The difference, though, is that they can be played with eye tracking, using hardware from Tobii and others that tracks your gaze as you look at the screen. You'll obviously need an eye tracker, but also a PC running the Windows 10 April 2018 Update or newer.



Digital Image Processing

Digital Image Processing means processing digital image by means of a digital computer. We can also say that it is a use of computer algorithms, in order to get enhanced image either to extract some useful information.

Image processing mainly include the following steps:

1.Importing the image via image acquisition tools;

2.Analysing and manipulating the image;

3.Output in which result can be altered image or a report which is based on analysing that image.

What is an image?

An image is defined as a two-dimensional function, F (x, y), where x and y are spatial coordinates, and the amplitude of F at any pair of coordinates (x,y) is called the intensity of that image at that point. When x, y, and amplitude values of F are finite, we call it a digital image.

In other words, an image can be defined by a two-dimensional array specifically arranged in rows and columns.

Digital Image is composed of a finite number of elements, each of which elements have a value at a particular location. These elements are referred to as picture elements, image elements, and pixels. A Pixel is most widely used to denote the elements of a Digital Image.

Types of an image

BINARY IMAGE– The binary image as its name suggests, contain only two-pixel elements i.e. 0 & 1, where 0 refers to black and 1 refers to white. This image is also known as Monochrome.

BLACK AND WHITE IMAGE– The image which consist of only black and white color is called BLACK AND WHITE IMAGE.

8-bit COLOR FORMAT– It is the most famous image format. It has 256 different shades of colors in it and commonly known as Grayscale Image. In this format, 0 stands for Black, and 255 stands for white, and 127 stands for gray.

16-bit COLOR FORMAT– It is a color image format. It has 65,536 different colors in it. It is also known as High Color Format. In this format the distribution of color is not as same as Grayscale image. A 16-bit format is actually divided into three further formats which are Red, Green and Blue. That famous RGB format.

Image as a Matrix

As we know, images are represented in rows and columns we have the following syntax in which images are represented:

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & f(0,2) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & f(1,2) & \dots & f(1,N-1) \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ f(M-1,0) & f(M-1,1) & f(M-1,2) & \dots & f(M-1,N-1) \end{bmatrix}$$

The right side of this equation is digital image by definition. Every element of this matrix is called image element, picture element, or pixel.

DIGITAL IMAGE REPRESENTATION IN MATLAB:

$$f = \begin{bmatrix} f(1,1) & f(1,2) & \cdots & f(1,N) \\ f(2,1) & f(2,2) & \cdots & f(2,N) \\ \vdots & \vdots & & \vdots \\ f(M,1) & f(M,2) & \cdots & f(M,N) \end{bmatrix}$$

In MATLAB the start index is from 1 instead of 0. Therefore, f (1,1) = f (0,0).

henceforth the two representation of image are identical, except for the shift in origin.

In MATLAB, matrices are stored in a variable i.e X, x, input image, and so on. The variables must be a letter as same as other programing languages.

PHASES OF IMAGE PROCESSING:

1.**ACQUISITION–** It could be as simple as being given an image which is in digital form. The main work involves:

a) Scaling

b) Color conversion (RGB to Gray or vice-versa)

2.**IMAGE ENHANCEMENT–** It is amongst the simplest and most appealing in areas of Image Processing it is also used to extract some hidden details from an image and is subjective.

3.**IMAGE RESTORATION–** It also deals with appealing of an image, but it is objective (Restoration is based on mathematical or probabilistic model or image degradation).

4.**COLOR IMAGE PROCESSING–** It deals with pseudo color and full color image processing color models are applicable to digital image processing.

5.**WAVELETS AND MULTI-RESOLUTION PROCESSING–** It is foundation of representing images in various degrees.

6.**IMAGE COMPRESSION-**It involves in developing some functions to perform this operation. It mainly deals with image size or resolution.

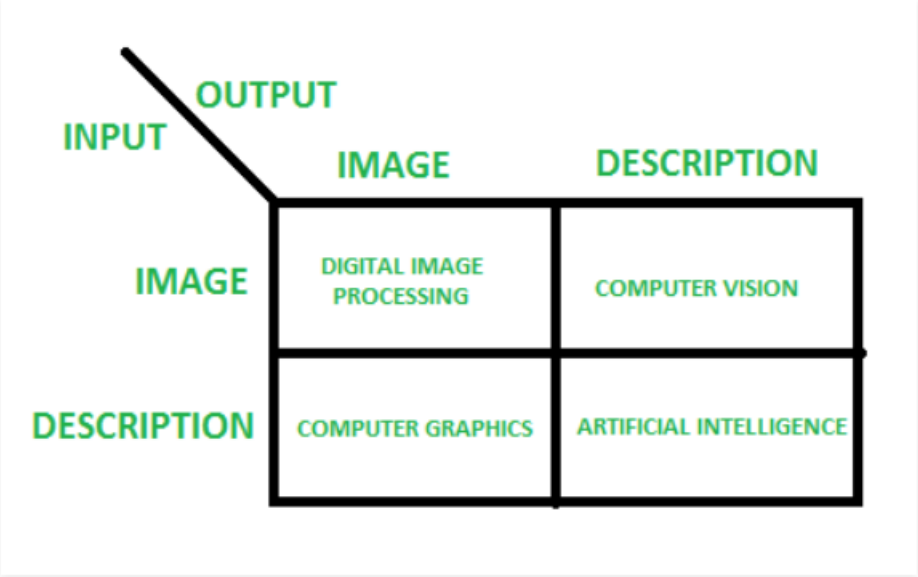
7.**MORPHOLOGICAL PROCESSING-**It deals with tools for extracting image components that are useful in the representation & description of shape.

8.**SEGMENTATION PROCEDURE-**It includes partitioning an image into its constituent parts or objects. Autonomous segmentation is the most difficult task in Image Processing.

9.**REPRESENTATION & DESCRIPTION-**It follows output of segmentation stage, choosing a representation is only the part of solution for transforming raw data into processed data.

10.**OBJECT DETECTION AND RECOGNITION-**It is a process that assigns a label to an object based on its descriptor.

OVERLAPPING FIELDS WITH IMAGE PROCESSING:



According to block 1, if input is an image and we get out image as an output, then it is termed as Digital Image Processing.

According to block 2, if input is an image and we get some kind of information or description as an output, then it is termed as Computer Vision.

According to block 3, if input is some description or code and we get image as an output, then it is termed as Computer Graphics.

According to block 4, if input is description or some keywords or some code and we get description or some keywords as an output, then it is termed as Artificial Intelligence.

References:

- **https://www.microsoft.com/en-us/research/product/eyes-first/**
- **https://www.tobii.com/tech/technology/what-is-eye-tracking/**
- **https://en.wikipedia.org/wiki/Digital_image_processing**
- **https://en.wikipedia.org/wiki/Pattern_recognition**
- **https://www.geeksforgeeks.org/digital-image-processing-basics/**