

基於事件相機眼動追蹤之動態虛像生成抬頭顯示系統設計

梁瑋倫¹

¹ 工業技術研究院/電子與光電系統研究所

Email : liangweilun@itri.org.tw

摘要

相較於傳統影像感測器，事件相機（Event camera）具備極高的時間解析度與低延遲特性，特別適合追蹤如眼球般快速移動的目標，且其亦可輕鬆應對於任何周圍環境亮度下觀察動態影像的任務。因此，事件相機在眼動追蹤（Eye tracking）的應用上，相較於傳統數位相機更具效能。本文提出一種整合事件相機與抬頭顯示器（Head-up display, HUD）的智慧顯示系統，透過即時眼動資料，於駕駛注視之車外實體物體鄰近處動態生成虛像資訊。該系統可使駕駛在行駛過程中，可於原視線方向快速獲取車輛相關資訊，無須將視線轉移至固定虛像顯示位置，如同傳統 HUD 所需。此設計有助於降低視覺注意力轉移所造成的認知負荷，進而提升駕駛專注力與行車安全性。

1、簡介

隨著智慧車載技術的快速發展，駕駛人與顯示介面之間的互動效率，已成為提升行車安全與操作便利性的重要關鍵。抬頭顯示器(Head-up display, HUD)[1]可提供駕駛在行車過程中，無需低頭查看儀表板，即可於擋風玻璃前方所生成的虛像[2]獲取諸如車速、速限、導航、時間等的重要資訊。駕駛可藉此避免於行駛途中，得反覆低頭查看儀表板，而導致開車過程中無法全神集中注意力的情形。

然而傳統的抬頭顯示器僅會於擋風玻璃前方固定的位置生成虛像，當汽車在行駛的過程中，若欲觀看抬頭顯示器所產生之虛像時，駕駛必須將目光從觀看車外實體物的情況中，切換至抬頭顯示器所生成之虛像的位置，再將目光切換回觀看車外實體物。如此來回切換目光，除了容易造成駕駛頭暈目眩，亦無法將注意力始終集中於觀看車外之實體場景，仍可能造成短暫的注意力分散與認知負荷增加，進而影響行車安全。若能設計出於觀看前方景象時亦可同時於鄰近前方景象的位置處生成虛像的光學系統，則可大幅提升駕駛眼睛的舒適度以及行車上的安全。

相對於傳統數位相機而言，事件相機(Event camera)[3]以較高的速度記錄著亮度的變化狀態，從而具有較高的時間解析度與動態變化感知能力，而得以捕捉到更快速的動作。詳細地說，事件相機除了僅需回傳簡單的亮度變化資訊外，各個感光單元皆以異步編碼的形式記錄著像素的亮度變化，由於所有像素紀錄事件的時間戳均為獨立、並可互不相同，各別獨立的感光單元只要一感測到亮度上的變化，便可即時進行記錄，故此具有低時間延遲的特性。而無須像傳統相機得於連續相等的時間間距中，費時蒐集一張張完整的圖片。相較於傳統逐幀式相機，事件相機能以更低延遲、更高頻率回應眼動變化，尤其適合應用於駕駛等高速視覺任務場景。

傳統相機若使用全域快門(Global shutter)[4]的工作模式，得在同一時刻將所有的感光單元同時曝光，並耗時彙整於此同一時間點所曝光的所有感光單元使其構成一張完整的圖片，才能再次蒐集下一時刻的圖片影像。若是使用滾動式快門(Rolling shutter)[4, 5]的工作模式，則不是一次打開所有感光像素進行曝光，而是按照行或列的順序逐行或逐列掃描進行曝光，並於固定的時間間距中，獲取一張完整的圖片。

在執行眼動追蹤(Eye tracking)[6]的任務時，由於眼珠可高速地運動，故適合藉由具備高時間解析率之事件相機來捕捉眼珠中顏色較深之角膜和瞳孔的位置及其移動的方向。此外，由於事件相機具備能於高亮及深暗環境中捕捉動態影像的特性，因此無論在白天或黑夜，皆可有效地利用事件相機來觀察眼珠的運動軌跡。綜合事件相機的優異特性，本文提出一個先藉由事件相機追蹤駕駛目光所觀看的方位，再藉由抬頭顯示器於鄰近駕駛觀看車外實體物的位置處產生虛像的光學系統。使得駕駛在開車的過程中，可同時觀看車外的實體場景以及鄰近實體物的位置所生成的虛像。藉此來避免駕駛者在使用傳統抬頭顯示器時，必須將目光於車外實體場景以及虛像兩者間進行長距離切換的過程，進而提升注意力集中與駕駛安全性。

2、事件相機

事件相機，又稱動態視覺感測器(Dynamic vision sensor, DVS)，是一種受生物視覺啟發的仿生視覺感測器，其工作方式完全不同於以全域快門或滾動式快門的方式來進行工作的傳統相機。全域快門指的是所有的感光單元在曝光時會同時曝光，並不存在時間上的差異。滾動式快門則是指在曝光時逐行讀取相機感測器上的像素值。兩者皆是在一個固定時間間段中，獲取一張完整的圖片後，才能再次於下一個時間間段來獲取另一張圖片。

然而對於事件相機中的單一像素僅會獨立地輸出所接收之光強度所產生的變化資訊，一旦所處於某個位置之像素的亮度值發生改變並達到一定的變化量時，事件相機便會回傳一個用以闡述亮度變化狀態的事件。當亮度由昏暗變為明亮時，以 On event 來進行記錄；反之，當亮度由明亮變為昏暗時，則以 Off event 來進行記錄；而當亮度維持不變時，則以 No event 來進行記錄。藉由集合位於各個不同位置的像素所偵測到的亮度變化資訊，即可記錄場景在不同時刻的變化狀態。

圖 1(a)描述一台事件相機與一台傳統相機針對相同物體進行攝影的過程。被攝影的物體中，左側為一個往右移動的實心圓，右側則為一個靜止的實心三角形，而攝像的結果則分別呈現於圖 1(b)與圖 1(c)。見圖 1(c)，假設用於攝影的傳統相機，為一台具有一百萬像素之傳統數位相機，其影格速率(Frame rate)為 10 fps，而 $10 \text{ frame/second} = 1 \text{ frame}/(0.1 \text{ second})$ ，亦即從第 0 秒起，由影像感應器所接收到的一張具有一百萬像素之畫面影像，必須花費 0.1 秒鐘的時間，將光信號接收後，轉換為電信號，再進行傳遞、分析與儲存等一系列的處理流程，最終形成一張完整的數位影像後，才能再接收下一張具有一百萬像素之畫面影像。因而於第 0 至第 0.1 秒鐘之間的時段，將無法再接收並處理其他張畫面影像。

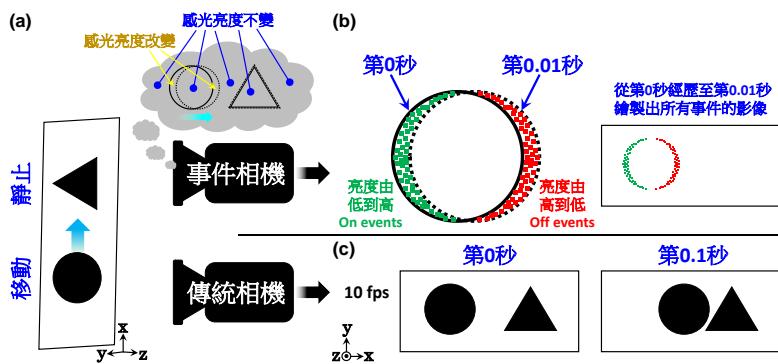


圖 1 事件相機的原理

傳統數位相機所拍攝到的影像是以色光三原色(紅色-R、綠色-G、藍色-B)合成各像素點的顏色。一旦感測器接收到外界光信號後，便會針對光信號進行分析，將紅、綠、藍三原色的色光中，較強的光信號以較大的數值來進行標記。反之，較微弱的光信號則以較小的數值來進行標記。因此，RGB 可使用三個引數來進行表示，引數的數值介於 0~255 之間，藉由不同的比例相加混色，即可混合出各式各樣的顏色。

假設當一個感光像素接依序收到三個不同狀態的外界光訊號，第 1 個狀態為：淺粉紅(顏色最淺、最明亮)、第 2 個狀態為：天空藍(顏色較淺粉紅更深)、第 3 個狀態為：深棕色(顏色較天空藍更深)。傳統數位相機必須將可合成各個顏色狀態的 RGB 亮度值進行分析與紀錄，故此所需處理的數值資料相對較為複雜，這對硬體在處理上的負擔也會相對沉重。然而事件相機在這三種不同外界光訊號的變化過程中，處理上則僅會記錄其亮度處於提升或降低的狀態。亦即當淺粉紅轉變為天空藍時，對於事件相機而言，其亮度屬於由高到低之間的變化，以 Off event 來進行記錄；而當天空藍轉變為深棕色時，對於事件相機而言，亦屬於亮度由高到低之間的變化，故此也以 Off event 來進行記錄。由於傳統數位相機必須藉由記錄 RGB 亮度絕對值，來合成出各種複雜的顏色狀態。反觀事件相機僅會記錄外界光訊號的亮度為提升、降低或不變，進而大幅降低所需處理之信息的複雜度，由於可避免大量數據的傳輸，其能耗將遠低於傳統相機。

此外，由於事件相機中的各個感光單元乃是以異步編碼的形式來記錄像素之亮度所產生的變化，因此並不存在傳統相機“幀率”(Frame per second, FPS)，也就是一秒鐘的影片含有多少張靜態圖片的概念。傳統相機在獲得一張靜態圖片時，乃是藉由全域快門或滾動式快門的形式，將所有的感光單元在一個固定時間段中進行曝光，並彙整所有的像素值來形成一張完整的圖片。然而由於事件相機中的各個像素僅會獨立地輸出所接收之光線的強度變化資訊，故此無需耗時等待如同傳統相機的「曝光」時機，而得以藉由各獨立像素高速地記錄亮度的變化狀態來提升時間的解析率，進而捕捉到更快速的動作。

由於事件相機中的每個像素乃是獨立地紀錄亮度變化的狀態，因而得以捕捉到極亮和極暗區域之亮度所產生的變化。而不會如同傳統相機般，因著得在一個時間段中獲取一張完整的圖片，而得顧及整張攝影圖片的亮度，當亮度超過或低於了應有的曝光範圍時，會受到過曝或欠曝的限制。

事件相機中，每個像素的光感測器乃為獨立運作，而非如同傳統相機般，得在一個時間段中，藉由將所有感光像素進行曝光的方式來取得一張完整的圖片。因此，事件相機只要有一個像素所偵測到的亮度值產生變化，便會回傳一個事件。譬如圖 1(a)，藉由事件相機觀測一個黑色實心圓(低亮度)在白色背景(高亮度)上的運動，而黑色實心三角形則處於靜止的狀態。並將觀測的結果呈現於圖 1(b)，前一時刻與下

一時刻的位置分別以實線及虛線進行繪製。

當黑色圓於事件相機光感測器上的成像以水平(x 軸方向)位移一個像素的距離時，在同一垂直高度上(相同 y 軸座標)之水平排列(不同 x 軸座標)的一列像素中(假設此列像素穿過黑色圓於光感測器上所成之像的中心)。黑色圓於此列像素上所成的像之左側輪廓將從其中一個像素向右位移至另一個像素上，而僅造成一個像素由原先被黑色圓成像覆蓋的低亮度狀態，變成被白色背景成像所照亮的高亮度狀態，進而呈現 On event 的狀態；同理，黑色圓於此列像素上所成的像之右側輪廓將從其中一個像素向右位移至另一個像素上，而僅造成一個像素由原先被白色背景成像所照亮的高亮度狀態，變成被黑色圓成像覆蓋的低亮度狀態，進而呈現 Off event 的狀態。如此，此一水平排列且通過黑色圓成像中心的像素中，被黑色圓成像的左側輪廓向右划過的像素，將逐一由 No event 轉變為 On event，再變回 No event 的狀態；而被黑色圓成像的右側輪廓向右划過的像素，則逐一由 No event 轉變為 Off event，再變回 No event 的狀態。

只要有一個感光像素所偵測到的亮度值產生變化，感光像素隨即便會記錄亮度的變化資訊。因此，若觀測由像素陣列所構成之整體影像感測器，在黑色實心圓成像移動的過程中，僅會於圓形輪廓處產生藉由少量的 On 與 Off events 所連繫成的一個具有極其細微邊框的空心圓圖樣(左半圓輪廓生成 On events、右半圓輪廓生成 Off events)。此外，圖 1(a)中的三角形由於始終處於一個靜止的狀態，因而不會產生事件流[見圖 1(b)]。

當黑色實心圓成像於事件相機之影像感測器上移動的過程中，由於事件相機儲存事件之時間間隔可以小於 1 微秒，假設黑色實心圓所移動的速度，造成事件相機於此時時間間隔中所記錄的 On 與 Off events，恰好可以構成一個具有極其細微邊框的空心圓圖樣。然而本文於解說的過程中，為了有利讀者觀看，在此刻意將事件相機儲存事件之前後時刻的間隔進行提升，並設置成 0.01 秒，以至於當黑色實心圓在移動的過程中，前一時刻與下一時刻間之位移的距離較大，進而繪製出較多的 On 與 Off events[見圖 1(b)]。

3、抬頭顯示器搭配眼動追蹤技術

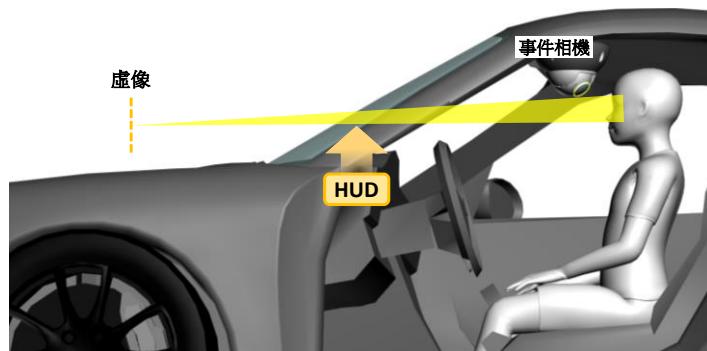


圖 2 藉由事件相機進行眼動追蹤後，再藉由抬頭顯示器生成適合駕駛觀看之虛像位置

眼珠的運動可達到相當高的移動速度。由於事件相機具有高時間解析率的特性，因此適合先使用事件相機追蹤汽車駕駛的雙眼所觀看的方向，再藉由抬頭顯示器將虛像產生於駕駛目光所觀看的位置(見圖 2)。使得汽車在行駛的過程中，駕駛若欲獲得抬頭顯示器所提供的資訊，可直接觀看鄰近於目光所關注之車外實體物的位置所生成的虛像。而無須如同傳統抬頭顯示器般，在觀看虛像的過程中，得先將觀看車外實體場景之目光挪移至固定位置處所生成的虛像，再返回至觀看車外的實體場景。因此可避免目光在不同位置上來回切換的過程中，駕駛會產生分心與昏眩的情形。

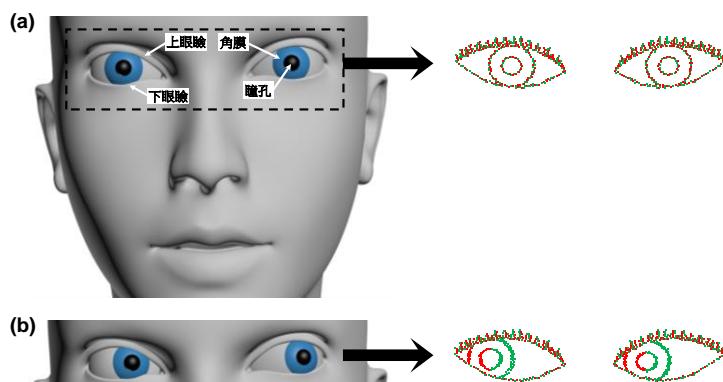


圖 3 藉由事件相機針對眼睛區域進行攝影，並將攝影的結果呈現於右方

圖 3 呈現藉由事件相機進行眼動追蹤的範例，左側為被攝影的眼睛狀態，右側為藉由事件相機所拍攝出的影像。在此刻意將事件相機所捕捉到的影像進行水平鏡像翻轉，以便對應駕駛朝汽車擋風玻璃前方所觀看的視角。在圖 3(a)中，即便長期將目光停留在固定的位置，瞳孔、角膜與上下眼瞼依舊會因著汽車引擎運作所導致的震動而生成均勻分佈於輪廓處的 On 與 Off events。在圖 3(b)中，由於角膜的顏色較眼白更深，而瞳孔的顏色又較角膜更深的緣故，當目光往左方位移的瞬間，事件相機所捕捉瞳孔與角膜的影像，將於此兩者輪廓的左側呈現 Off events、右側呈現 On events。

3.1 目光朝左方、前方、右方觀看

圖 4 乃藉由事件相機執行眼動追蹤，來捕捉眼珠中顏色較深之角膜和瞳孔的位置及其移動的軌跡(上方圖)，進而搭配抬頭顯示器，於適合眼睛觀看的位置來生成虛像(下方圖)。見圖 4(a)，當目光往左方位移的瞬間，事件相機所捕捉瞳孔與角膜的輪廓影像，將於左側呈現 Off events、右側呈現 On events。此時抬頭顯示器應當於擋風玻璃左方生成虛像，以提供適合駕駛之目光所觀看的方位。見圖 4(b)，當目光定睛於前方時，由於瞳孔與角膜保持於固定的位置，因此事件相機僅產生因著汽車引擎運作所導致的震動而生成均勻分佈於瞳孔、角膜與上下眼瞼輪廓處的 On 與 Off events。此時可分析雙眼之瞳孔與角膜位於上下眼瞼中心的位置，來判斷出此時目光正朝著前方觀看，進而要求抬頭顯示器於擋風玻璃中間(駕駛座位的前方)生成虛像。見圖 4(c)，當目光往右方位移的瞬間，事件相機所捕捉瞳孔與角膜的輪廓影像，將於右側呈現 Off events、左側呈現 On events。此時抬頭顯示器應當於擋風玻璃右方生成虛像。

在圖 4(a)與圖 4(c)中，若長期將目光停留在往左或往右觀看的位置，瞳孔與角膜亦會因著保持於固定位置的緣故，而使得事件相機僅會產生因著汽車引擎運作所導致的震動而生成均勻分佈於瞳孔、角膜與上下眼瞼輪廓處的 On 與 Off events。此時可分析雙眼之瞳孔與角膜是位於上下眼瞼左側還是右側的位置，來判斷出此時目光是朝著左側還是右側觀看，進而要求抬頭顯示器於擋風玻璃左方或右方生成虛像，以提供適合駕駛之目光所觀看的方位。

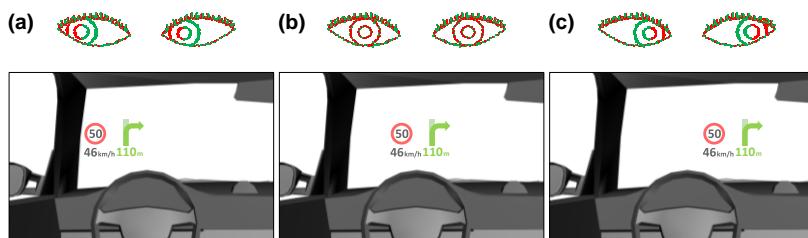


圖 4 事件相機捕捉目光所觀看的方向 (a) 目光往左方位移的瞬間 (b) 目光持續向右觀看 (c) 目光往右方位移的瞬間

3.2 目光觀看遠處或近處之物體

當前方空間較為寬敞、無障礙物時，目光可望見前方較遠處的物體。然而當駕駛欲觀測前方較鄰近汽車的物體時，目光則需要聚焦在較鄰近的位置。譬如當前方突然有行人經過時，為要從觀看遠處場景改為觀看近處行人，目光便會從遠處切換成聚焦在前方較為鄰近駕駛的位置。

倘若當目光關注前方不同遠近之物體時，能將 HUD 所生成的虛像位置隨著目光所關注的前後位置一同產生變化，將可避免駕駛在觀看遠近距離固定之虛像與遠近距離不同的實體物時，需長距離地前後切換目光所關注的位置。若欲生成不同遠近位置之虛像，可採用雙焦面抬頭顯示器[7]或可連續變焦的抬頭顯示器[8]，在擋風玻璃前方任意可變的前後位置產生位於遠處或位於近處的虛像。

見圖 5(a)，當眼睛望向遠方的樹木時，虛像也應當被抬頭顯示器投射於較遠處的位置，使得眼睛在觀測遠方的物體時，可同時聚焦在實體物(樹木)與虛像的位置上。即便光學系統在設計上無法使得虛像全然落於與實體物具有相同的出光面，也可以將虛像調整至較遠的位置，使得駕駛當目光於實體物與虛像二者間切換時，較不會因著長距離的切換而感到疲憊。然而當前方突然有行人經過時[如圖 5(b)]，目光則需聚焦在前方較鄰近駕駛的位置處(也就是行人的身上)，此時若抬頭顯示器將虛像維持在如圖 5(a)較遠處的位置，則無法讓駕駛看清虛像，因此可將虛像的位置調整至鄰近行人的位置，避免在觀察行人時，若欲觀看虛像，還需長距離地將目光切換至較遠處的虛像位置。

見圖 5(a)，當被觀測物位置距離駕駛較遠時，光線發散至雙眼的角度(聚散角，Vergence angle)較小，從而使得雙眼之瞳孔間距離較大。反之，見圖 5(b)，當被觀測物位置距離駕駛較近時，光線發散至雙眼的角度則較大，從而導致雙眼之瞳孔有著較短的間距，而呈現出類似鬥雞眼的形象。

圖 5(b)右下角之圖呈現當雙眼觀看近處物體的瞬間，事件相機所捕捉到的影像。此時雙眼之瞳孔與角膜於靠近鼻子的部分呈現 Off events，至於靠近耳朵的部分則呈現 On events。因此，當事件相機接收到此種影像時，即可判斷左右眼瞳孔與角膜正處於相互靠攏的階段，便可要求抬頭顯示器將所生成的虛像往靠近駕駛的位置移動。

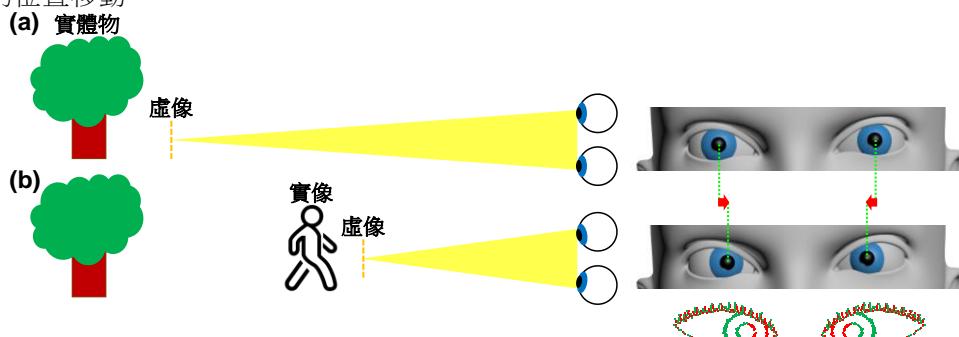


圖 5 藉由抬頭顯示器生成遠近不同之虛像，搭配觀測觀測不同距離之實體物

若目光長期停留在觀看近處物體的狀態，瞳孔與角膜則會因著保持於固定位置的緣故，而僅產生因著汽車引擎運作所導致的震動所生成均勻分佈於瞳孔、角膜與上下眼瞼輪廓處的 On 與 Off events。此時則可藉由觀測雙眼瞳孔間距較平時為短，以及雙眼外觀處於鬥雞眼型態來研判目光正處於觀看近處物體的狀態。同理，當雙眼從觀看近處物體轉為觀看遠處物體或定睛停留在長時間觀看遠處物體時，亦可藉由同樣的方式來進行分析。

藉由上述法則，當駕駛在觀測前後不同遠近的實體物時，可藉由事件相機快速地捕捉雙眼瞳孔在移

動過程中的影像。當雙眼瞳孔相互遠離、並呈現較長的瞳孔間距時，抬頭顯示器應當生成距離駕駛較遠的虛像。反之，當雙眼瞳孔彼此靠攏、並呈現較短的瞳孔間距時，則生成距離駕駛較近的虛像。

4、結論

由於眼珠可高速地進行運動，適合藉由具備高時間解析率之事件相機作為眼動追蹤的工具，來捕捉駕駛的目光。且為因應行車安全，於夜間駕駛時應當關閉車內燈源。而事件相機恰好有著可於極亮或極暗環境中觀測動態影像的能力，所以無論白天或黑夜皆可藉由事件相機有效地執行眼動追蹤的任務。憑藉事件相機的這些特性，本文提出一種結合事件相機與動態虛像生成技術的智慧型 HUD 系統。透過即時追蹤駕駛的注視位置，系統能根據駕駛視線所關注之車外實體物體，動態生成資訊虛像，實現「視線所及資訊所示」的顯示模式，無須額外移動視線至特定 HUD 位置，進而提升注意力集中與駕駛安全性。

我們利用事件相機來捕捉眼珠中顏色較深之角膜和瞳孔的位置及其移動的方向，進而研判其目光所觀看的位置。藉此讓駕駛在觀看車外實體物時，可直接於鄰近被觀看之實體物的位置處生成虛像，駕駛無須將目光於車外實體場景及某個固定位置處的虛像兩者間進行長距離的切換，從而避免駕駛得分心抽出目光觀看如同傳統抬頭顯示器所生成之固定位置的虛像，而造成無法始終集中注意力觀看車外實體場景的情形，且可避免目光於切換過程中所產生的疲勞感，進而提升行車安全。本系統預期可應用於智慧交通、增強現實輔助駕駛（AR driving assistance）等場域，具有實用性與技術前瞻性，為未來人因導向與視覺對齊式顯示技術提供實證依據與設計參考。

5、參考文獻

- [1] Z. Lv, J. Liu and L. Xu, "A Multi-Plane Augmented Reality Head-Up Display System Based on Volume Holographic Optical Elements With Large Area," *IEEE Photonics J.*, vol. 13, pp. 1-8, Oct. 2021.
- [2] H. Hong, J. Park and H. Shin, "Increase of Virtual Image Area by Reflection of Mirrors," *J. Disp. Technol.*, vol. 3, pp. 421-425, Dec. 2007.
- [3] I. Schiopu and R. C. Bilcu, "Lossless Compression of Event Camera Frames," *IEEE Signal Process. Lett.*, vol. 29, pp. 1779-1783, Aug. 2022.
- [4] B. Fan, Z. Wan, B. Shi, C. Xu and Y. Dai, "Unified Video Reconstruction for Rolling Shutter and Global Shutter Cameras," *IEEE Trans. Image Process.*, vol. 33, pp. 6821-6835, Nov. 2024.
- [5] A. Otsuka, Z. Tang, S. Lu and T. Yamazato, "Code Synchronization in Visible Light Communication System Using Dual Orthogonal Rolling Shutter Image Sensors," *IEICE Commun. Express*, vol. 13, pp. 450-453, Dec. 2024.
- [6] X. Zhang, S. -M. Yuan, M. -D. Chen and X. Liu, "A Complete System for Analysis of Video Lecture Based on Eye Tracking," *IEEE Access*, vol. 6, pp. 49056-49066, Aug. 2018.
- [7] Z. Qin, S.-M. Lin, K.-T. Luo, C.-H. Chen, and Y.-P. Huang, "Dual-Focal-Plane Augmented Reality Head-Up Display using a Single Picture Generation Unit and a Single Freeform Mirror," *Appl. Opt.*, vol. 58, pp. 5366–5374, Jul. 2019.
C.-T. Mu, W.-T. Lin, and C.-H. Chen, "Zoomable Head-Up Display with the Integration of Holographic and Geometrical Imaging," *Opt. Exp*