

À propos de l'écran de réalité virtuelle du scaphandre ummite

Julien Geffray

27 octobre 2020



En octobre 2020, le physicien Mark Brongersma (université de Stanford) en partenariat avec des chercheurs de Samsung annonce la publication d'un papier dans une revue scientifique* résumant les recherches ayant abouti à la création de prototypes d'écrans offrant des résolutions (densité de pixels) extrêmement élevées et jamais atteintes auparavant : de 10 000 à 20 000 ppp† ! Des densités à comparer à celles des *smartphones* modernes, dont les résolutions déjà considérées comme élevées tournent autour de 300 à 500 ppp. Ces recherches visent à produire des écrans ultra haute définition pour les solutions VR (réalité virtuelle) tels que casques et lunettes, où l'*effet de grille* ou capacité de l'œil à distinguer les pixels individuellement est pour le moment très présent et limite le rendu réaliste de ces dispositifs binoculaires.

<https://news.stanford.edu/?p=36817>

Cet article m'a donné l'idée de calculer la résolution de l'écran de réalité virtuelle du scaphandre EEWE ANIXOO à partir des données indiquées dans la D69 reçue en juin 1968, il y a plus de cinquante ans.

* Won-Jae Joo, Mark L. Brongersma et al. (23 oct 2020) "Metasurface-driven OLED displays beyond 10,000 pixels per inch", *Science* Vol. 370, n° 6515, pp. 459-463. DOI: [10.1126/science.abc8530](https://doi.org/10.1126/science.abc8530)

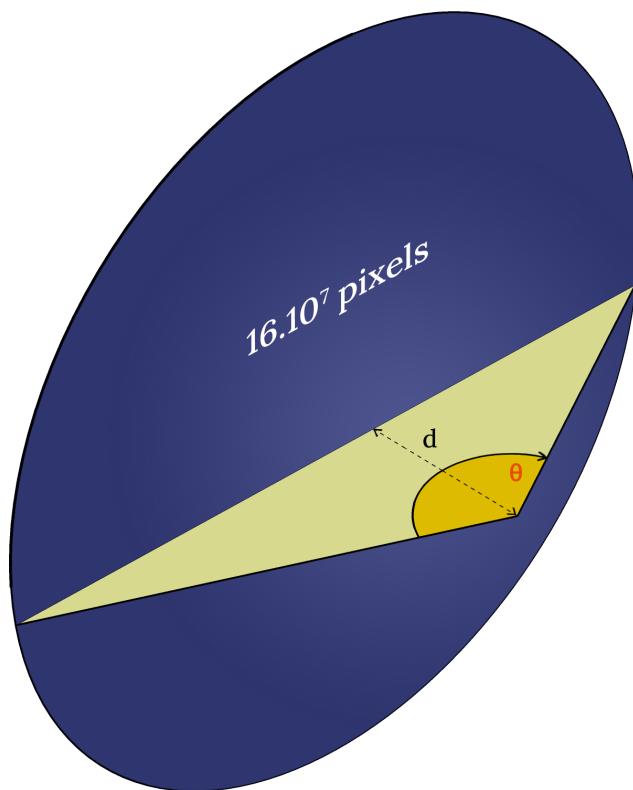
† points ou pixels par pouce, en anglais dpi (*dots per inch*, utilisé en impression) ou ppi (*pixels per inch*, utilisé pour les écrans). C'est une densité *linéique*, un nombre de points sur une distance de 1 pouce (soit 2,54 cm).

À la note n° 6 de ce document, il est indiqué (passages qui nous intéressent mis en gras) :

Le EEWEANIXOO constitue ce que vous appelleriez un "scaphandre" ou combinaison. [...]

L'enceinte qui sépare l'épiderme du EEWE subit, **au niveau du visage, un élargissement considérable de forme tronconique**. La base d'un tel tronc de cône est située depuis l'œil, et selon un angle de **130° sexagésimaux**, à une **distance de 23 centimètres**. Elle représente un écran pourvu sur sa surface de quelque **16.10⁷ centres excitables** capables de diffuser chacun, avec divers niveaux d'intensité, tout le spectre électromagnétique compris entre 3,9.10¹⁴ et 7,98.10¹⁴ cycles/seconde*. **La définition des images obtenues est suffisamment élevée pour que les deux yeux ne puissent faire la différence entre les perceptions visuelles normales et celles engendrées artificiellement par cet organe.**

Supposons l'écran circulaire et schématisons :



avec $d = 23 \text{ cm}$

$$\text{et } \theta = 130^\circ \times \frac{\pi}{180} = 2,26893... \text{ radians}$$

la surface de l'écran en bleu est égale à :

$$S = \pi \left(d \tan \frac{\theta}{2} \right)^2 \approx 7\,643 \text{ cm}^2$$

Par ailleurs, son diamètre est de 98,65 cm. Pratiquement un mètre !

La définition de l'écran étant de 160 millions de pixels†, la densité de celui-ci (nombre de pixels par unité de longueur) ou résolution est égale à :

$$\text{res} = \sqrt{\frac{\text{def}}{S}} = 144,68 \text{ pixels/cm}$$

ou, exprimé plus conventionnellement en pixels par pouce (1 pouce = 2,54 cm) :

$$\text{res} = 367,5 \text{ ppp}$$

Soit à peine plus que la résolution de l'iPhone 4 (326 ppp) commercialisé en 2010, *smartphone* dont Apple qualifiait l'écran de "Retina". Mais est-ce comparable ?

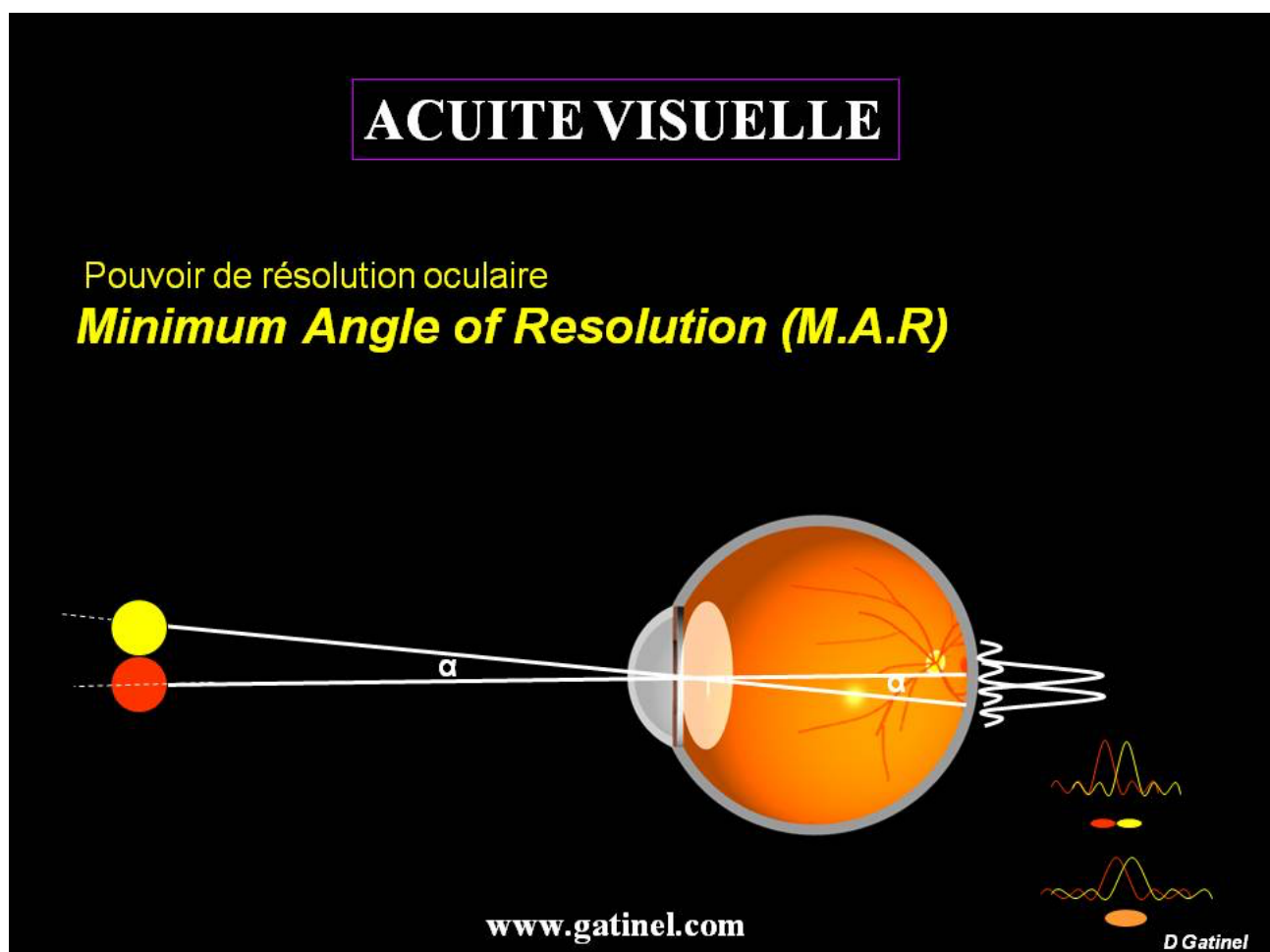
* Soit des longueurs d'onde de 375,5 à 768,5 nanomètres, ce qui recouvre effectivement le spectre visible de la vision humaine.

† C'est encore davantage qu'un écran 16K (dont aucun n'est commercialisé à l'époque de la rédaction de cet article) qui offre une définition de 15 360 × 8 640, soit plus de 130 millions de pixels.

Retina est un terme marketing signifiant que l'œil ne perçoit plus les pixels individuellement. Il est fonction de la densité de pixels de l'écran, mais aussi de la distance à laquelle cet écran est observé en pratique. Si on approche davantage l'écran des yeux, la résolution doit être plus élevée pour rester "Retina". Si l'écran est au contraire conçu pour être regardé à une distance plus éloignée (un téléviseur par exemple) cette résolution peut être plus basse et l'écran restera tout de même "Retina".

Il est plus simple d'utiliser le terme "Retina" que "reste en dessous du pouvoir séparateur de l'œil pour telle distance". Mais comment calcule-t-on cette densité limite en fonction de la distance ?

Soit s l'écart entre deux objets lointains (ou plutôt ici, deux points consécutifs sur un écran proche) et d la distance de ces deux points à l'iris de l'œil. Il existe un angle de vision α minimum, en dessous duquel l'œil humain sera incapable de discerner les deux points individuellement et n'en verra plus qu'un seul. Cet *angle minimum de résolution* définit le *minimum séparable* d'acuité visuelle*.



Pour un être humain disposant d'une bonne vision, cet angle α vaut 1 minute d'arc, soit $1/60^\circ$ de degré†. Pour des personnes à la vision un peu moins bonne – ce qui est courant – cet angle peut monter jusqu'à 5 minute d'arc ($1/12^\circ$ de degré) voire au delà pour les gens ayant une mauvaise vue. On considère la valeur optimale de 1 minute d'arc.

* Voir par exemple <https://www.gatinel.com/recherche-formation/acuite-visuelle-definition/mar-angle-minimum-de-resolution/> dont est tirée l'image ci-dessus (© Dr Gatinel).

† Les constructeur d'écrans parlent de la limite de "60 pixels par degré de la fovéa". L'angle minimum de $1/60^\circ$ de degré permet effectivement de caser 60 pixels par degré.

On a la relation :

$$\tan \frac{\alpha}{2} = \frac{s}{2d}$$

En réarrangeant, la distance minimale entre deux points est :

$$s = 2d \tan \frac{\alpha}{2}$$

En fait, l'angle α étant extrêmement petit, les conditions de Gauss sont atteintes et l'approximation des petits angles est utilisable de telle sorte que :

$$\tan \alpha \approx \alpha = \frac{s}{d}$$

L'écran de l'EEWE étant situé à $d = 23$ cm, en considérant la limite de 1 minute d'arc de pouvoir séparateur, la distance théorique maximale d'un pixel à l'autre pour cet écran est :

$$s = 0,0067 \text{ mm (soit 67 microns)}$$

Qu'en est-il en pratique dans le texte ? On obtient, avec les caractéristiques décrites dans le D-69 :

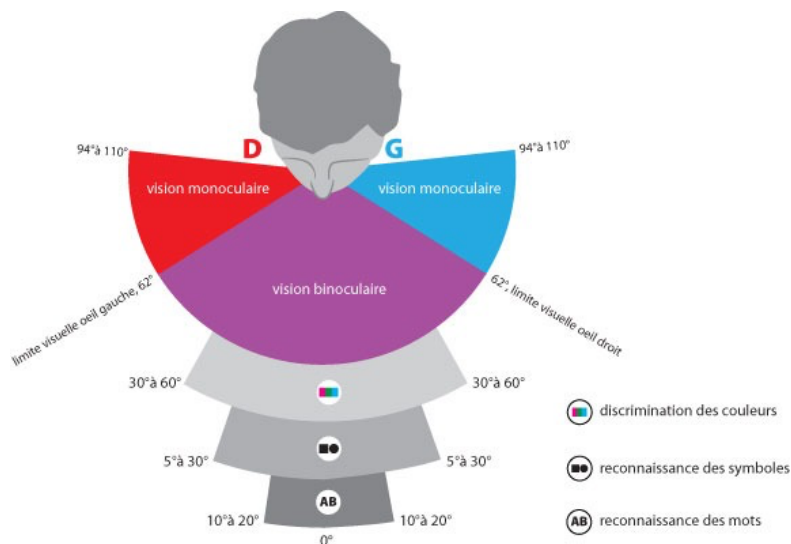
$$s = \frac{1}{\text{res}} = 0,0069 \text{ mm (soit 69 microns)}$$

Par conséquent, en épluchant et en moulinant les quelques paramètres disponibles décrivant l'écran du EEWEANIXOO comme un appareil binoculaire dont la résolution est *“suffisamment élevée pour que les deux yeux ne puissent faire la différence entre les perceptions visuelles normales et celles engendrées artificiellement pour cet organe”*, on obtient effectivement – à 2 microns près, ce qui est remarquable – exactement les caractéristiques optiques permettant de pouvoir donner à cet écran décrit en 1968 le qualificatif moderne de “Retina”. Le ou les individus qui ont rédigé ce texte n'ont donc pas mis ces chiffres au hasard.

Certes, ils s'agit de simple trigonométrie et d'optique géométrique basique. Mais dans les années soixante, imaginer un casque de réalité virtuelle à une époque où il n'était même pas envisagé d'atteindre ce genre de résolution pour des imprimantes à encre sur papier reste un bel exercice, qui en outre, vu depuis l'accomplissement technologique de notre époque un demi-siècle plus tard, a le mérite de rester cohérent.

Une note sur l'angle de vision de 130° : la vision "complète" humaine est plus large : 190° à 220° . Mais elle n'est binoculaire que sur environ 125°. Le reste sur les côtés est flou et monoculaire. C'est sans doute la raison pour laquelle l'écran binoculaire de l'EEWE couvre un angle solide de 130° et pas davantage.

À propos de la vision binoculaire stéréoscopique, en relief ou “3D”, celle-ci est assurée en envoyant deux images différentes à chaque œil. Comment est-ce possible sans passer par deux écrans proches distincts séparés par une cloison verticale au niveau du nez ; ou sans porter de lunettes passives polarisées ou de lunettes actives ?



Le texte de la D69 donne la réponse :

La **vision binoculaire** est réalisée grâce à la disposition prismatique de chaque centre émetteur. L'excitation des faces opposées, de sorte qu'un œil n'ait pas accès à l'image ou mosaïque vue par l'autre, est réalisée d'une manière très complexe (image A*).

Un transducteur enregistre les champs électriques engendrés par les muscles oculaires des deux globes (véritables électromyogrammes). Le XANMOO connaît ainsi, à chaque instant, l'orientation de l'axe de la pupille. D'autres part, les prismes excitables qui composent l'écran (ceux-ci étant de dimensions microscopiques) sont situés sur la surface d'une couche d'émulsion visqueuse qui leur permet de tourner librement. Ces prismes sont contrôlés mécaniquement par un double champ magnétique, de sorte que la moitié d'entre eux obéissent à une composante horizontale du champ, et l'autre moitié à la transversale. De cette manière, l'un et l'autre groupe orientant leurs faces indépendamment, comme ces stores vénitiens utilisés par les Terriens qui orientent indépendamment leurs lames quand on tire sur leurs cordelettes, afin de régler l'angle d'incidence de la lumière (dans ce cas les "cordes" seront les deux champs magnétiques, et le facteur moteur la réponse du XANMOO aux micro-mouvements musculaires du globe oculaire).

Il s'agit donc d'un système asservissant l'orientation des "lames en store vénitien" des micro-prismes orientables en direction de chaque œil, à une technologie d'*eye tracking*... à ceci près que cette dernière n'est pas fondée comme on en a l'habitude désormais sur un suivi optique automatisé de l'iris en mouvement imagé par une caméra, mais directement par les mesures électriques des muscles de l'œil.

Revenons pour terminer à l'annonce de l'université de Stanford. Avec des résolutions annoncées de 10 000 à 20 000 ppp, l'écran de l'EEWE semble-t-il dépassé ? À vrai dire, il ne s'agit pas du tout des mêmes solutions technologiques, ce qui fait que les résolutions ne sont pas directement comparables.

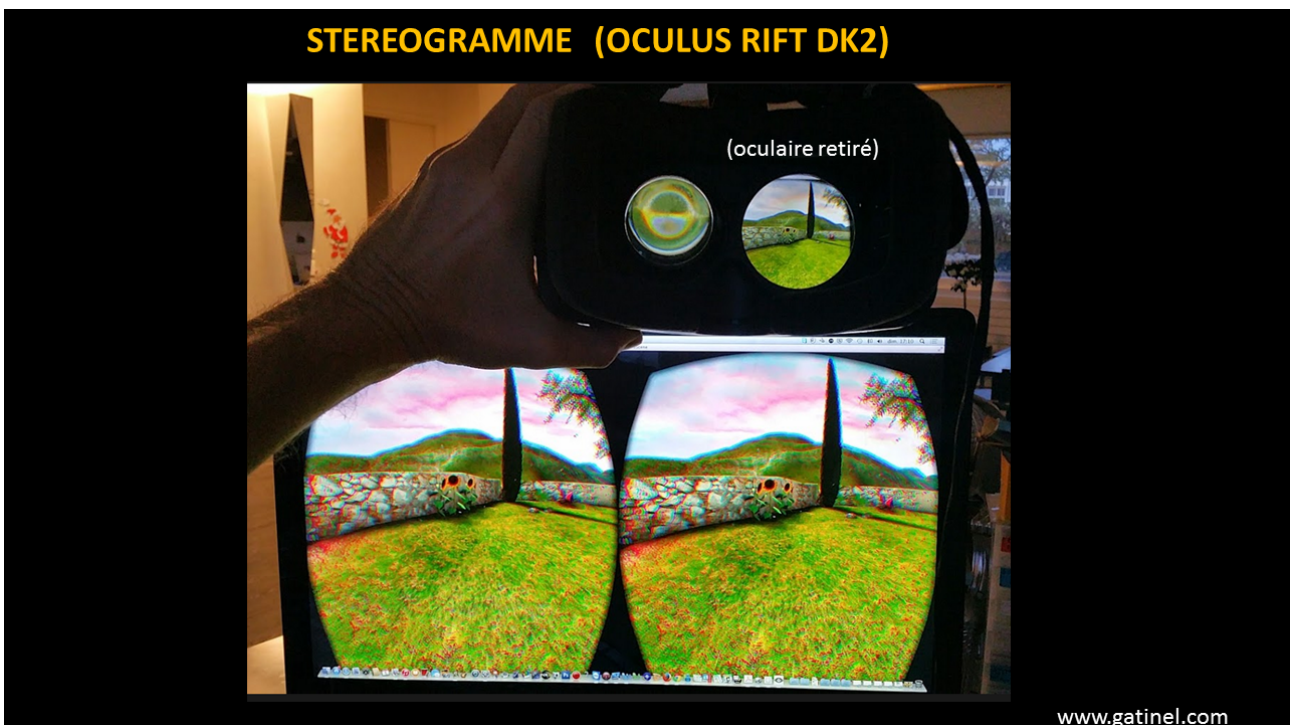
L'écran du EEWE est "en champ libre" et à bonne distance. Il n'y a rien qui perturbe le trajet des rayons lumineux jusqu'aux yeux. Au contraire, les écrans des casques VR habituels sont comme des masques de plongée dont on aurait remplacé chacun des deux verres séparés par un écran spécifique, de part et d'autre de la cloison nasale.

* Cette image est malheureusement manquante dans l'original.



Plongez dans le monde merveilleux de la réalité virtuelle

Or, comme ces deux écrans sont très proches des yeux, il est exclu de pouvoir accommoder la vision directement. Il est nécessaire d'interposer un oculaire muni de lentilles qui focalisent les rayons lumineux sur la fovéa, au fond de l'œil.



*Une vue plus réaliste. L'oculaire de droite a été retiré et on aperçoit à la place l'image de l'écran.
En dessous, un moniteur de bureau affiche les deux images projetées par chaque écran du casque,
dont l'observation séparée par chaque œil recrée une vision binoculaire artificielle en relief.*

En procédant ainsi, l'image virtuelle est en réalité projetée beaucoup plus loin des yeux que ne le sont les deux écrans physiquement. Autre conséquence : les lentilles ont pour effet de zoomer dans l'écran. Voici sans doute les deux raisons pour lesquelles on cherche à propulser la densité de pixels sur les casques VR à des niveaux faramineux bien plus élevés que sur les *smartphones* :

- les écrans sont trop proches des yeux.
- les lentilles grossissent les pixels.

Ajoutons que les lentilles introduisent aussi toutes sortes de distorsions. La solution technique du EEWEANIXOO qui en est dépourvue et qui nécessite une densité de pixels bien moindre est-elle plus judicieuse techniquement ? On notera tout de même que le nombre total de pixels (160 millions) à gérer dans le grand écran du EEWE semble particulièrement énorme, même en le comparant à de futurs casques VR dont les (petits) écrans seraient boostés à 10 000 voire 20 000 ppp, et qui mettront pourtant au défi les puissances de calcul disponibles.