

Projet de position relative aux conditions de déploiement des caméras dites « intelligentes » ou « augmentées » dans les espaces publics

**Commentaires du CNPEN
2022-03-18**

Avez-vous des commentaires et suggestions sur la partie 1 (« observations préalables ») ?

Avant d'aborder la partie 1, une observation sur le titre et le vocabulaire. Pour un tel texte, le vocabulaire doit être précis et non ambigu. C'est aussi l'occasion de mettre en œuvre une terminologie claire et cherchant à éviter l'anthropomorphisation de dispositifs techniques. Il nous semble par conséquent souhaitable d'éviter l'utilisation de qualificatif tels qu' « intelligent » ou « augmenté » pour les dispositifs techniques objets de ce projet. Nous faisons des propositions dans ce sens plus avant.

Dans ce contexte, il nous semble important de définir précisément les fonctions des différents composants techniques concernés par le texte.

Une caméra vidéo est un dispositif permettant d'obtenir via des capteurs d'images et de sons un signal vidéo éventuellement couplé à un signal audio. Noter qu'en particulier le signal vidéo est le résultat de traitements de signaux élaborés permettant d'obtenir une bonne qualité d'image (balance des blancs, mise au point, équilibrage des couleurs, etc.). Sauf exception, une caméra embarque donc par construction des capacités de traitements d'image élaborés.

Une caméra vidéo est attachée à un dispositif fixe (poteau, mur) ou mobile (personne, véhicule, drone) en ayant éventuellement par rapport à son dispositif d'attache certains degrés d'orientation dynamique.

Le signal vidéo produit par une caméra peut-être soit enregistré dans la caméra soit transmis via un réseau filaire ou radio à d'autres dispositifs. Ces dispositifs peuvent consister en éléments de mémorisation, de visualisation ou de traitement numérique.

Les vidéos captées peuvent être utilisées en temps quasi réel ou en temps différé (de quelques secondes à un nombre indéterminé années). Ces utilisations résultent d'analyses conduites par des opérateurs humains qui peuvent être aidés par des mécanismes automatisés de diagnostic produits à partir des vidéos et possiblement croisées avec d'autres sources d'information (heure de la journée, météo, température, etc.)

Dans ce contexte c'est l'opérateur humain dont la compétence est augmentée par les capacités de traitement automatisé des vidéos plutôt que les caméras elles-mêmes.

Il nous semble donc souhaitable de parler de Système d'analyse automatisée de vidéo (SAAV) plutôt que de vidéo « augmentée », et par ailleurs de n'utiliser les termes « vidéo-protection » ou « vidéo-surveillance » que dans des contextes bien spécifiques et identifiés. Par exemple « vidéo-protection » induit d'idée de protection (de quoi, comment, pourquoi, avec quels moyens ?), associant au dispositif une finalité a priori positive qui n'a rien d'avéré.

(1.3) Dans ce contexte le paragraphe exclut d'utiliser les SAAV à des fins de reconnaissance biométrique. Or l'identification biométrique s'inscrit dans le continuum des dispositifs de vidéos considérés ici, ce pourquoi nous pensons qu'il serait intéressant de l'inclure dans le spectre du texte. En outre, il s'agirait également de mentionner le fait que l'identification n'est pas uniquement

biométrique et ainsi souligner que ce n'est pas parce que le projet ne porte pas sur l'identification biométrique qu'il n'aborde pas d'autres techniques d'identification.

(1.14) « ...définir collectivement ... » : comment ? est-ce l'enjeu du projet ?

Par ailleurs, l'expression « ligne rouge » qui mériterait d'être remplacée par : définir quels dispositifs, ou quels usages de tels dispositifs, devraient être proscrits ou limités [dans nos démocraties].

Avez-vous des commentaires et suggestions sur la partie 2 (« La vidéo « augmentée » : portrait d'une technologie aux multiples usages ») ?

L'utilisation des termes « vidéoprotection » ou « vidéosurveillance » doit faire l'objet d'une attention particulière. En effet elle induit une finalité (i.e. protection ou surveillance) à un dispositif de captation vidéo.

Il faut être attentif à bien identifier et qualifier chacun des maillons de la chaîne d'élaboration des informations depuis la captation jusqu'à interprétation des signaux vidéo.

(2.1.2) Le terme « probabiliste » n'est pas ici approprié et il semblerait plus correct de parler de statistique ou « à partir de modèles à base de connaissances ».

Il serait en outre important de définir ce que signifie, ici, caméra vidéo, mais également ce que recouvrent plus précisément les espaces publics (2.1.3).

(2.1.4) il faudrait remplacer « Même si ces algorithmes s'intègrent à des caméras vidéo... » par « Même lorsque ces algorithmes s'intègrent à des caméras vidéo... »

(2.2.3) ensemble d'exemples intéressants : noter que dans les attributs cités, le sexe et l'âge sont des indications biométriques.

Le cas de l'industrie est intéressant cependant, ici, le projet concerne les espaces publics dont il nous semble qu'une usine ne fait pas partie.

(2.2.4) «... auparavant humaine... » ça le restera certainement car les SAAV aident la plupart du temps les opérateurs humains à prendre des décisions (qu'on espère de la sorte plus informées).

(2.2.5) L'exemple « détection d'émotions » suppose des capacités de reconnaissance faciale.

Il faudrait remplacer « sous une forme identifiante » par « sous une forme pouvant permettre l'identification », cette première formulation semblant sous-entendre que des éléments d'identification sont présents dans l'image captée par le dispositif (p.ex. le nom marqué sur le visage). Le terme d'« anonymisation » semble également inapproprié car nous en connaissons la quasi impossibilité aujourd'hui.

(2.2.8) « l'automatisation permettant d'économiser les coûts salariaux des opérateurs de vidéosurveillance » : (1) « permettant » ou « pouvant permettre » ? (2) citer la source de l'affirmation, l'analyse des gains ici est supposée avérée, et cela n'a rien d'évident.

(2.3.2) éviter des phrases peu précises comme « des flux vidéo basées sur des algorithmes d'« intelligence artificielle ». » : « le développement de technologies d'analyse automatique des flux vidéo » suffirait.

(2.3.4-5) Quelles références et sources ?

Avez-vous des commentaires et suggestions sur la partie 3 (« Une technologie porteuse de risques gradués pour les droits et libertés des personnes ») ?

Dans cette partie, il est important de définir - et même plus généralement dans le projet - ce que sont les données personnelles et les données sensibles. De plus, les caméras ne sont pas « intelligentes », elles ne constituent que des dispositifs de captation vidéo.

Une caméra peut capter (et combiner) deux types de signaux : la vidéo et l'audio, il convient donc de préciser si ces caméras captent uniquement l'image ou également le son.

(3.1.1) « l'image des personnes est visionnée » -> elle est surtout analysée.

(3.1.2) Les opérateurs humains peuvent également capter voire analyser ou prendre en compte les données sensibles des personnes, donc le risque est également présent dans les systèmes de surveillance vidéo classiques.

(3.1.4) Ce ne sont pas les personnes qui sont analysées, mais les vidéos ; puis on infère (correctement ou pas, dans une coopération entre le SAAV et les opérateurs humains, qui ne semble pas bien prise en compte dans l'ensemble du projet) de ces vidéos des éléments concernant les personnes. Cette inférence est liée à l'interprétation humaine (en général) du diagnostic produit par le SAAV. Le terme « probabiliste » est encore présent ici de manière inappropriée à notre sens.

(3.1.5) Il faudrait remplacer le terme de « vidéoprotection » par l'expression « captation vidéo ». Paragraphe important qui demanderait à être détaillé. La phrase « permettraient de résoudre par eux-mêmes » s'appliquant à des dispositifs devrait être remplacée par « permettraient de résoudre automatiquement »

(3.1.6) Redondance avec 3.1.4 et utilisation d'un exemple concernant l'analyse du visage d'une personne et donc de la reconnaissance faciale (cf. 1.3).

(3.1.7) Le sens du paragraphe est intéressant et utile. La "préservation de l'anonymat dans l'espace public" ici invoquée, est-elle un droit, une liberté ? Il faut noter que l'analyse d'éléments visuels n'est en rien nouvelle et qu'elle est historiquement faite par des humains (passants, voisins, agents de la force publique, etc.), et continuera dans la plupart des cas à l'être, aidée (?) par des SAAV.

(3.1.9) « ...ils n'ont pas de moyen d'avoir conscience que celles-ci peuvent, non pas seulement les filmer, mais également les analyser » : ce texte devrait les aider, non ? le constat n'est donc pas définitif et éducation, information et expérience devraient aider.

(3.1.10) Excellente et très importante remarque. Notez que l'exemple final contredit (1.3).

(3.1.12) semble à étayer ! De plus « intrusivité » devrait être défini, le terme n'existant pas en français.

(3.2) Il faudrait expliciter comment l'usage est effectué.

Sur les conditions de déploiement, il s'agirait de préciser par qui ces dispositifs sont déployés, pour quel usage et sous quel contrôle / comment.

Il s'agit donc de définir ce que signifie le déploiement (de quoi, par qui, pour quoi et comment). Cet exemple montre que le SAAV peut-être fortement décorrélé de l'ensemble des prises d'images.

(3.2.5) Il serait bon de préciser la notion de vulnérabilité utilisée ici.

Avez-vous des commentaires et suggestions sur la partie 4 (« Des conditions de légalité différenciées en fonction des objectifs, des conditions de mise en œuvre et des risques des dispositifs de vidéo « augmentée ») ?

Globalement dans cette partie importante visant à déboucher sur des propositions de réglementation ou de normalisation et sans entrer dans le détail des mesures proposées, il nous semble :

1. Important de prendre en compte les remarques terminologiques évoquées en introduction de nos commentaires. Par exemple, il faudrait remplacer des expressions comme « risques des dispositifs de vidéo « augmentée » » par « risques induits par les usages des systèmes d'analyse automatisée de vidéo ».
2. Qu'en conséquence, il est effectivement difficile pour le CSI de statuer sur des dispositifs de vidéo « augmentée » alors que leur définition n'est pas explicite (4.1.3)
3. Que la rédaction doit être plus précise : p.ex. dans 4.2.2 les images ne contiennent pas des personnes (mais des images pouvant permettre l'identification de personnes)
4. Que les applications médicales potentielles ne semblent pas être prises en compte (4.2.5.3)
5. Qu'il conviendrait de préciser que cela ne peut pas « être légalement mis en œuvre » au regard de la Loi Française (4.2.5.4) voir Européenne
6. Que la notion de transparence – importante – devrait être précisément introduite (4.2.7.2) et illustrée
7. Que les exemples (4.4.3.3) et (4.4.3.4.) sont éclairants et montrent que dans (4.4.3.5) l'usage des techniques automatisées doit être pris en compte.
8. Que la notion de consentement qui se pose globalement ici doit faire l'objet d'une réflexion générale : on la retrouve de manière multiple dans les enjeux d'éthique du numérique par exemple dans le cadre de l'internet des objets (voir le rapport récent de France Stratégie).

Finalement, nous pensons qu'une analyse des termes du titre du projet de position doit être effectuée relativement à la terminologie et aux positions proposées.

Notons enfin que le CNPEN s'est auto-saisi sur les enjeux d'éthique relatifs aux technologies de reconnaissance faciale, posturale et comportementale et a pour objectif de proposer des recommandations concernant les usages de ces technologies.