Présentation générale du cours
Traitement des images et géométrie

Cnam - UV 18458

Ph. Destuynder

Octobre 2003

Nous commençons par définir l'objet de notre investigation : l'image numérique. Puis après avoir discuté quelques aspects technologiques des capteurs d'images numériques (appareils photos et caméscopes numériques), nous proposons une approche intuitive des problématiques qui sont traitées dans les chapitres suivants. En particulier nous essaierons de définir une sorte de cahier des charges des solutions à trouver.

Les images numériques.

Précisons l'objet Mathématique de nos réflexions pendant tout ce cours.

Une image numérique est une fonction notée f de deux ou trois variables spatiales. Elle est à valeurs dans $\mathbb{R}^n$, où n est égal à 1 pour des images monochromatiques et à 3 pour des images couleurs. Cette fonction sera toujours supposée de carré intégrable sur l'ouvert de visualisation sur lequel elle est définie appelé l'écran en bidimensionnel et le volume visuel en tridimensionnel. Cet ouvert sera un rectangle (en bidimensionnel) ou un parallèlipipède (en tridimensionnel). Nous le noterons Ω dans toute la suite. La quasi totalité du texte traite d'images bidimensionnelles .

En fait, dans la pratique, la fonction f qui caractérise l'image n'est pas définie en tous les points de l'ouvert Ω mais seulement sur une grille de points appelés *pixels* en bidimensionnel (*picture element en anglais*), et *voxels* en tridimensionnel (*volume element* toujours en anglais). Ceci est dû à un échantillonage de l'image rendu nécessaire par les prises de vue numériques (on ne peut stocker qu'une quantité finie d'informations). Bien entendu plus il y a de points, plus le volume de stockage de l'image est important. Par exemple une image télévision standard nécessite en bidimensionnel), 512x682 = 349.525 pixels. Pour une séquence de 10 secondes il faut donc (25 images secondes) :

10x25x349.525 = 87.381.333 pixels. Si chaque valeur de la fonction f est stockée sur 8 bytes (soit 64 bits), il faut presque 1 gigabytes de mémoire pour stocker 10 secondes d'images au standard télévision dont la netteté n'est pas la propriété la plus évidente. Or un DVD*[1] peut stocker environ 5 Gigabytes soit environ une minute au format précédent et pourtant on y loge plus d'une heure de film ayant ce format de description. Il y a donc un secret ! C'est le compactage . Mais il faudra être très astucieux pour réduire d'un facteur d'au moins 60 la quantité d'informations à stocker. Notons également que l'image définie par ses valeurs aux pixels ou voxels, s'apparente à un tableau de chiffres que nous appelerons l'image informatique . Bien entendu il y a un lien étroit entre les deux types d'images évoquées ci-dessus et que nous expliciterons dans le premier Chapitre.

Le procédé de numérisation d'une image.

Un rayon optique contient a priori un mécanisme ondulatoire contenant un large spectre de fréquences. Chacune d'elles correspond à une couleur. En utilisant des filtres optiques ou polariseurs, on peut extraire d'un rayon la composante Rouge, la Bleue ou la Verte. Les rayons qui sortent des polariseurs s'apparentent à des lumières monochromatiques. Ils viennent percuter des cellules photoélectriques qui génèrent une tension électrique fonction croissante de l'intensité lumineuse. Mais une cellule ne voit qu'une couleur. La synthèse des couleurs se fait ensuite en recombinant trois cellules photoélectriques successives. Le stockage de l'image numérique se fait donc à partir de l'ensemble des tensions enregistrée sur les cellules.

Les appareils de photos numériques du commerce possèdent entre 2 et 12 millions de pixels. Mais les caméscopes en ont beaucoup moins. Lorsque l'on veut reproduire une information tridimensionnelle, par exemple en utilisant un procédé stéréoscopique, il est nécessaire de comprendre comment la différence de chemins optiques entre deux rayons, permet de retrouver un indicateur de profondeur. Un peu de géométrie est alors nécessaire. Il faut ensuite intégrer un décodage des informations dans le système de stockage des deux images en mémorisant les positions des cellules qui permettent de les appareiller deux à deux.

[1] Digital Video Disc

La nécessité de corriger des images, les réparations.

Il y a plusieurs raisons qui peuvent conduire l'utilisateur à vouloir intervenir sur une image. Bien entendu, on pense en premier lieu à la réparation de l'image suite à une détérioration par exemple. Mais aussi, pour créer des effets spéciaux, mettre en valeur certaines parties ou au contraire, en faire disparaître d'autres. De façon plus standard, un agrandissement met souvent en évidence des contours en escalier qui sont peu réalistes. C'est le cas pour les images satellites ou celles provenant de matériels d'observation médicale. Un lissage permet de retouver des courbes plus régulières mais souvent au prix d'un étalement donnant un effet de flou. Il devient alors nécessaire de redessiner les lignes de contour et d'améliorer la netteté. Il s'agit là d'un challenge essentiel dans la procédure de traitement d'images numériques.

La reconnaissance et l'extraction d'objets.

Pour des raisons de moyens technologiques ou plus classiquement à la suite de *zoom* sucessifs, la définition d'une image peut être très insuffisante au point de ne pas permettre la reconnaissance des objets qu'elle contient. Certain d'entre nous se souviennent peut-être avoir passé des heures à retrouver les formes d'un foetus dans une échographie par ultrasons que seul l'amour parental permet de deviner. Les frères Schlumberger ont été les premiers à utiliser les problèmes inverse d'échographie pour détecter la présence de nappes de pétrole dans les sous-sols. Ce *sport* est devenu l'un des plus importants de l'industrie pétrolière et occupe des milliers d'ingénieurs dans le monde entier. Le principe est d'éclairer (acoustiquement parlalnt), un objet dans l'espace et d'observer l'écho qu'il nous renvoit. L'éclairement est réalisé par un mécanisme ondulatoire. Les ultrasons en sont un. En outre, ils présentent l'avantage d'avoir de faibles longueurs d'ondes et ainsi de permettre l'identification de petits objets sans perturber le milieu dans lequel ils se propagent. Donnons un exemple plus difficile : celui de l'analyse de sillages derrière un véhicule. Un écoulement d'air est (en général) invisible. Néanmoins il se créé des nappes tourbillonaires sur lesquelles les ondes sonores peuvent se réfléchir. Leurs échos permettent une identification de leur présence. On comprend alors comment cette démarche de l'analyse d'images, utilisant des modèles mathématiques, permet d'extrapoler les données insuffisantes pour reconstituer une réalité physique. Mais bien entendu il est alors nécessaire de disposer d'un modèle représentant cette physique. Les principes d'une telle démarche : remplacer l'ignorance des données par une connaissance de la physique sous-jacente, est une activité qui connaît de grands développements dans les industries médicales, pétrolières, mécaniques et bien d'autres.

Les mondes tridimensionnels réel et virtuel.

Quand Walt Dysney a démarré la création de dessins animés, il est peu probable qu'il ait entrevu le devenir extraordinaire de cette discipline rendu possible grâce à l'ordinateur. Poutant, les bases du dessins permettant de représenter les objets en tridimensionnels datent de G. Monge et de l'Ecole de Metz. La géométrie descriptive qui est le point de départ de toute représentation mathématique et donc informatisable, de scènes tridimensionnelles est devenue une discipline moderne après avoir fait souffrir tant de génération d'élèves sur leur planche à dessin. En fait, c'est le phénomène de la vision binoculaire qui est le pendant physiologique de la "descro" et par conséquent de la vision stéréoscopique. Le couplage de ces deux aspects (physiologique et mathématique) est le monde dans lequel évoluent les créateurs de jeux vidéos qui nous proposent des visualisations virtuelles de plus en plus réalistes. Nous abordons quelques aspects simples de cette activité dans ce cours.

Le codage et les corrections d'erreurs.

Les images, transmises par exemple par voie hertzienne, traversent des milieux perturbateurs comme des champs magnétiques. A cette occasion, des erreurs de transmission peuvent apparaître et détériorer l'image. Pour remédier à cette difficulté majeure, on multiplie l'information provenant du codage. Par exemple en ajoutant des bits de contrôle comme le bit de parité, (si la somme des bits de l'information est pair le bit de parité vaut 0 et 1 dans le cas contraire). En fait on peut aussi résumer les techniques de codage/décodage et de correction d'erreurs de la façon suivante. Le chiffre à coder est noté x. C'est par exemple un élement du corps $(Z/2Z)^n$. On lui applique une transformation linéaire injective notée C qui est a valeurs dans $(Z/2Z)^p$ avec $p > n$ et qui a pour mission d'éloigner les points images deux à deux. Ce qui signifie qu'il y aura $p - n$ bits de contrôle. Un des énormes avantage du corps utilisé est qu'il contient un nombre fini d'éléments. Pour cette raison la distance entre deux éléments reste finie. On utilise alors un des résultats fondamentaux de l'analyse matricielle qui permet de décomposer l'espace $(Z/2Z)^p$ en d'une part l'image de $(Z/2Z)^n$ par C et d'autre part le noyau de l'application transposée ${}^tC$ qui va de $(Z/2Z)^p$ dans $(Z/2Z)^n$. L'information codée est l'élément y de $(Z/2Z)^p$. Si elle est dans l'image de C on retrouve x. Sinon on la corrige en projettant y sur cette image à l'aide de la matrice $[{}^tC.C]^{-1}.{}^tC$. Bien entendu la matrice C est connue par l'émetteur et le récepteur. Il existe ensuite des façons plus ou moins astucieuses (et rapides!) de construire C. Ces aspects qui connaissent de nombreux développements sont abordés dans ce cours.