

# EXAMEN NSY116 - Multimédia et Interaction humain-machine - Février 2006

Durée : 2h - Tous documents autorisés

## CORRIGE INDICATIF

### Question 1. (2 points)

Le domaine du multimédia a été présenté en introduction du cours par 4 aspects. Définir ce que sous tendent chacun de ces points en quelques lignes : **(a)** convergence des médias **(b)** interactivité **(c)** hypermédia **(d)** immersion.

*remarque : on ne traite ici que du multimédia et pas de l'informatique en général*

*a) convergence : avec la multiplication des médias (pour l'écrit, l'image et le son) est apparue progressivement l'idée d'information, indépendante du support qui la véhicule. Le développement de l'informatique a ensuite permis de trouver dans la numérisation un formalisme commun pour manipuler plus facilement cette information, en particulier pour la transmission sans erreur (telos) la compression (stenos) et le secret (cryptos)*

*b) interactivité : les technologies multimédia permettent de produire des œuvres dont le contenu dépend non plus seulement de leur auteur, mais aussi des actions (des choix) du récepteur. Les dispositifs d'interaction humain-machine utilisés à ce jour reposent essentiellement sur la métaphore WIMP (Windows, Icon, Menu, pointer) mise au point au début des années 70.*

*c) hypermédia : généralisation aux informations non textuelles (image, son) du principe de l'hypertexte (Ted Nelson, années 60) qui consiste à décrire les associations fixes ou potentielles entre divers documents (liens). Le Web n'est que partiellement hypermédia (pas d'hypervidéo)*

*d) immersion : solliciter tous les sens (vision, audition, toucher, odorat, goût) du récepteur de l'œuvre pour lui en faciliter la compréhension (ex. des simulateurs de vols ou des jeux) ou au contraire de le surprendre (en jouant de paradoxes perceptifs). L'immersion est réalisée par une simulation (visuelle et sonore, le plus souvent) de l'environnement du récepteur.*

### Question 2. (2 points)

**(a)** A quoi sert la lemmatisation d'un corpus ? **(b)** Donner un exemple de logiciel qui la réalise.

*a) Lemmatiser un corpus consiste à regrouper tous les mots d'une même racine (par ex. chanter, chantons, chantonner, chant, chanson => chan). Cela permet ensuite de faciliter les recherches dans un corpus textuel. Exemple vu en cours : interrogation sur le corpus du journal Le Monde. Avant lemmatisation, seuls 22% des mots sont communs aux années 1992 et 93, après lemmatisation, le taux monte à 66%. La lemmatisation du français est difficile car de nombreux mots sont homographes (été, fût, etc.). Pour l'anglais, il existe en revanche un algorithme très efficace (algo. de Martin Porter).*

*b) On utilise souvent la lemmatisation dans les moteurs de recherche (encyclopédies, pages jaunes)*

### Question 3. (4 points)

Ecrire l'algorithme d'un traitement d'image effectuant un lissage "gaussien". On supposera que l'image est déjà chargée en mémoire dans une matrice de pixels baptisée Image, de largeur L et de hauteur H.

```
-- on traite à part le case des bords (cf exercice binarisation dans l'ED Image)
-- pour le reste de l'image :
pour i:=2 à H-1
  pour j:=2 à L-1
    NImage [i,j] := Image[i-1,j-1] + 2*Image[i-1,j] + Image[i-1,j+1]
                  + 2*Image[i, j-1] + 4*Image[i,j] + 2*Image[i, j+1]
                  + Image[i+1,j-1] + 2*Image[i+1,j] + Image[i+1,j+1]
  fin pour
fin pour
Image := NImage -- pour recopier toute la nouvelle image dans l'ancienne
```

**Question 4. (2 points)**

- (a) En terme de traitement du signal, quelle est l'opération par le traitement vu à la question 3 ?  
(b) Existe-t'il un traitement similaire pour le son numérique ?

a) il s'agit d'un filtre passe-bas

b) oui, quand on souhaite par exemple retirer les hautes fréquences d'un son

**Question 5. (1 point)**

Existe-t'il un algorithme de compression avec perte utilisable à la fois pour des images fixes et du son numérique ? (on souhaite une réponse argumentée)

*Les algorithmes de compression avec perte sont spécifiques du multimédia. On utilise certaines caractéristiques des systèmes perceptifs de l'homme. Par exemple :*

*- pour l'image : les petites variations locales de chromaticité peuvent être retirées de l'image (JPEG)*

*- pour le son : l'effet de masque des basses fréquences intenses sur les hautes fréquences plus faibles présentes simultanément permet de retirer ces hautes fréquences (MP3)*

*A ce jour, il n'existe pas d'algorithme de compression avec perte utilisable à la fois pour le son et l'image fixe.*

**Question 6. (2 points)**

(a) Donnez et expliquez les différentes étapes du pipeline 3D standard.

**Triangulisation** : les différents polygones constituant les objets d'une scène sont découpés en triangles. Ceux-ci sont plus faciles à gérer : leurs sommets sont nécessairement coplanaires et ils sont plus faciles à colorier (ils sont convexes).

**Transformations géométriques** : changement de repère pour passer du repère global utilisé pour décrire la scène au repère de l'observateur (du point de vue).

**Test de visibilité** : détermination des triangles vus de dernière et élimination si l'on ne doit garder que les triangles vus de faces.

**Clipping** : rejet des triangles hors du volume de vue, découpage des triangles à cheval sur les frontières du volume de vue.

**Calcul des intensités lumineuses** : attribution d'une couleur aux sommets d'un triangle en fonction de son orientation par rapport aux sources lumineuses.

**Projection** : projection des sommets 3D sur la surface 2D de l'écran.

**Coloriage** : balayage du triangle pour attribuer une couleur à chacun de ses pixels.

(b) Quelles techniques permettent d'accélérer le rendu 3D temps réel.

*On a donné déjà deux techniques dans les étapes précédentes. Ce sont des techniques dites de culling qui permettent d'éliminer un certain nombre de triangles de la chaîne de rendu :*

- le test de visibilité ou backface culling,
- le clipping ou clôturage.

*S'ajoutent à ces techniques de culling des éléments logiciels (par exemple l'organisation spatiale et hiérarchique d'une scène dans un graphe de scène) ou matériels (par exemple l'accélération matérielle : prise en charge des fonctions du pipeline graphique par la carte graphique).*

**Question 7. (2 points)**

Classez ces termes dans les deux catégories « Lancé de rayon » et « projection de facettes » (un terme peut-être présent dans les deux catégories) : triangulisation, projection, intersection droite/géométrie, qualité, composition de scène, temp réel, Z-Buffer, interpolation des intensités lumineuses, interpolation des profondeurs, API 3D, jeux vidéos, films d'animation, POV, calcul des normales, texturage, opérations booléennes, interactivité.

| <b>Lancé de rayon</b>                | <b>Les deux</b>                | <b>Projection de facettes</b>                  |
|--------------------------------------|--------------------------------|--|
| <i>Intersection droite/géométrie</i> | <i>Composition de scène</i>    | <i>Triangulisation</i>                         |
| <i>Qualité</i>                       | <i>Texturage</i>               | <i>Projection</i>                              |
| <i>Films d'animation</i>             |                                | <i>Temps réel</i>                              |
| <i>POV</i>                           |                                | <i>Z-Buffer</i>                                |
| <i>Opérations booléennes</i>         | <i>(Opérations booléennes)</i> | <i>Interpolation des intensités lumineuses</i> |
|                                      |                                | <i>Interpolation des profondeurs</i>           |
|                                      |                                | <i>API 3D</i>                                  |
|                                      |                                | <i>Jeux vidéos</i>                             |
|                                      |                                | <i>Calcul des normales</i>                     |
|                                      |                                | <i>Interactivité</i>                           |

### Question 8. (3 points)

Nous souhaitons développer un jeu vidéo de course de voiture. Comme tout jeu récent, il doit contenir un certain nombre de fonctionnalités telles que :

- des menus de configuration, d'informations,
- de belles séquences de cinématique en début et en cours de jeu,
- la gestion de la musique et des effets sonores,
- une interaction fluide (temps réel) et une réponse réaliste de la voiture,
- une belle qualité d'images lors des phases de jeux : modèle des voitures, décors, effets graphiques ...

Quels techniques et technologies préconisez-vous ? Classez-les en trois catégories : composition de scène, effets visuels et sonores, programmation. Argumentez vos choix.

*Pour la composition de scène : le mieux est d'organiser les données dans un structure hiérarchique permettant d'optimiser le rendu. On pourrait utiliser un graphe de scène. En règle général, dans un jeu vidéo, on sépare le « monde » (ce qui est statique) des « atomiques » qui bougent. Pour le monde, on utilise souvent les BSP (une partition binaire de l'espace sous forme de secteurs). Pour les atomiques, une structure hiérarchique (graphe de scène) qui permet d'animer facilement des sous parties du modèles (exemple : les fichiers MD2 ou MD3 de quake). Ces différents aspects devront être programmés puisqu'ils ne sont pas intégrés dans les API 3D ou nous devons utiliser un middleware pour le jeu tel que Renderware, Ogre ou Irrlicht.*

*Pour les effets visuels et sonores : On peut utiliser DirectSound et DirectMusic pour les plateformes Windows/Xbox. Pour avoir accès à d'autres plateformes, on préférera Fmod ou OpenAL pour la gestion du son spatialisé et de la musique d'ambiance (au format MP3 ou WAV), possiblement accédés via SDL. Pour les effets visuels, DirectDraw ou SDL peuvent être utilisés selon les plateformes visées. Les cinématiques sont des vidéos qui peuvent être encodées en Divx ou plus généralement en Mpeg.*

*Pour la programmation : on a évoqué l'utilisation de SDL ou DirectX. Dans le premier cas, l'API 3D sera OpenGL et dans le deuxième Direct3D. Dans les deux cas, le langage le plus intuitif à utiliser pour programmer un jeu est un langage objet tel que C++ pour la performance ou Java pour la facilité et la richesse des bibliothèques.*

### Question 9. (2 points)

(a) Quels aspects de l'humain et de la machine limitent l'efficacité des systèmes interactifs actuels ?

*Du côté de l'humain, c'est notre capacité à percevoir qui est limitant par rapport à la possible richesse des médias (millions de couleurs avec des teintes très proches, fréquences sonores*

*inaudibles ou indiscernables, ...). Notre capacité à mémoriser, à trouver, à traiter l'information est également limitée comme nous avons pu le voir à travers « le processeur humain » proposé par Card et les mesures effectuées sur les capacités de traitement des différents éléments le constituant.*

*Du côté de la machine : on peut citer l'inadéquation entre l'interface graphique 2D et le nombre d'information à trier et à visualiser. Le couple clavier/souris ne semble pas non plus idéal pour interagir avec ces données de plus en plus nombreuses.*

**(b)** Pour la désignation directe de zones à l'écran, donnez une alternative à l'utilisation de la souris.

*La souris n'est clairement pas idéale pour désigner/sélectionner un élément à l'écran puisqu'elle nécessite un déplacement relatif à sa position actuelle. Un écran tactile offre donc cette possibilité de désigner directement un objet graphique par l'intermédiaire du doigt ou d'un stylet. On peut également envisager les techniques d'oculométrie qui permettent de déterminer quelle zone de l'écran est regardée par l'utilisateur.*