# Infographie tridimensionnelle

[](http://commons.wikimedia.org/wiki/File:Alexroomclose.jpg)

Image réaliste réalisée avec Maya et Mental Ray

La **synthèse d'image tridimensionnelles** souvent abrégée **3D** (3D pour trois dimensions : x,y,z, les trois axes qui constituent le repère orthonormé de la géométrie dans l'espace) est un ensemble de techniques notamment issues de la CAO qui permet la représentation d'objets en perspective sur un moniteur d'ordinateur. Elle est actuellement très utilisée en art numérique dans l'industrie du film, initiée par les studios Pixar et ILM et, depuis 1992, dans beaucoup de jeux vidéo. Il convient de ne pas faire la confusion avec les termes 3D relatifs au relief ou à la stéréoscopie.

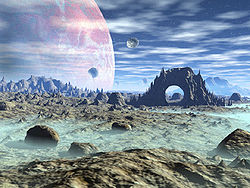
**Histoire**

La dénomination **images tridimensionnelles** est le nouveau nom donné à ce qu'on appelait dessin ou peinture en perspective depuis la Renaissance.  
Voici un exemple de vue perspective réalisée par Piero della Francesca en 1475[réf. nécessaire]. Certains peintres s'aident du perspectographe, *première machine à dessiner en 3D*, pour leurs compositions picturales.

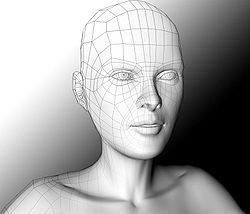
[](http://commons.wikimedia.org/wiki/File:Piero_della_Francesca_-_Ideal_City.jpg)

*La cité idéale* (1475), longtemps attribuée à Piero della Francesca

La différence est que, avant l'apparition des ordinateurs, la perspective était obtenue par des méthodes graphiques dérivées directement de la géométrie projective. Actuellement, les dessins sont calculés numériquement en partant des données numériques tridimensionnelles. Ceci permet de changer facilement le centre de projection et ses différents paramètres. En particulier, on peut calculer des séries de projections en déplaçant le point de projection et réaliser ainsi des animations.

[](http://commons.wikimedia.org/wiki/File:Bryce-demo-May00PicOfMonth.JPG)

Représentation en perspective d'un paysage, avec le logiciel Bryce 3D

[](http://commons.wikimedia.org/wiki/File:Polygon_face.jpg)

Rendu laissant apparaître la structure polygonale utilisée pour la synthèse de l'image.

[](http://commons.wikimedia.org/wiki/File:X-plane_artistic_render.jpg)

X-plane artistic render

Les images de synthèse informatiques actuelles reposent sur les mêmes principes de projection sur un plan et sont indissociables de l'histoire de l'informatique. Les images de synthèse ont débuté au début des années 1950, aux États-Unis, et étaient réservées à la recherche, notamment universitaire. On construisit un système composé d'un tube cathodique et d'un crayon optique, d'après une idée de Ivan Sutherland, du Massachusetts Institute of Technology (MIT), pour le contrôle aérien de l'armée de l'air, puis en 1961 on ajouta sur l'écran une croix pour indiquer la position du crayon optique. On pratiquait alors de l'image 2D, puis de l'image 3D, plus coûteuse en temps de calcul et financièrement.

Puis les universités se servirent également des images tridimensionnelles, et en 1967, l'université de l'Utah se spécialisa dans ce domaine, en particulier les professeurs David C. Evans et Ivan Sutherland, qui essayèrent de modéliser divers objets tels que la voiture d’Ivan Sutherland, et qui fonderont en 1968 la société Evans & Sutherland. Puis en 1970, Xerox créa le PARC (Palo Alto Research Center), qui travaillera très librement, car sans objectifs commerciaux ; il en sortira de nombreuses découvertes que Xerox ne saura pas exploiter. En 1975 fut créée une des plus célèbres images de l'infographie, la théière, devenue depuis un objet classique de test pour les applications 3D. La théière qui a servi de modèle repose maintenant au Boston Computer Museum, près d'un ordinateur qui reproduit son image en trois dimensions.  
Jusqu'aux années 1980, peu de personnes abordaient ce domaine en raison des coûts du matériel. Mais l'apparition des ordinateurs personnels, comme le Xerox Star, l'IBM-PC en 1981, et l'Apple Macintosh en 1984, ont démocratisé l'utilisation de la 3D pour l'étude et la production… L'évolution de la technologie aboutit à de superbes simulations de navettes ou de fusées par la NASA, ou de paysages et de visages.

Mais c'est à partir des années 1990 que l'image de synthèse et la 3D se démocratisèrent et se développèrent de façon importante, avec notamment l'arrivée de matériels plus puissants permettant le temps réel, comme des cartes 3D chez Silicon Graphics sur des stations de travail haut de gamme puis plus tard dans les ordinateurs grand public Amiga et PC avec des cartes 3D comme la 3DFX ou dans les consoles de jeux comme la PlayStation, la Dreamcast. Depuis, l'accélération 3D est partie intégrante des matériels informatiques.

**Généralités**

La synthèse d'image tridimensionnelles fait appel à un espace vectoriel. Cet espace est décomposé en trois dimensions sur les axes cartésiens, nommés habituellement *X*, *Y* et *Z*.

Prenons, dans une pièce, un point de référence, et définissons des directions gauche-droite *(X)*, avant-arrière *(Y)* et haut-bas *(Z)*. Pour aller de l'origine à un point donné, il faut faire :

* *x* mètres vers la droite ;
* *y* mètres vers l'avant ;
* *z* mètres vers le haut.

Si l'on peut permuter l'ordre dans lequel on peut faire le déplacement, cette combinaison (*x*, *y*, *z*) est unique : un point de la pièce est représenté par un triplet unique, et un triplet représente un seul point de la pièce (voir *Repérage dans le plan et dans l'espace*, *Système de coordonnées cartésiennes* et *Géométrie analytique*). Ce triplet de valeurs s'appelle les coordonnées du point.

Le point de référence est appelé « l'origine » et se note habituellement *O*, il correspond au triplet (0, 0, 0).

Se déplacer d'une valeur *a* à gauche équivaut à se déplacer d'une valeur −*a* à droite. Reculer d'une valeur *b* équivaut à avancer de −*b*. Descendre d'une valeur *c* équivaut à monter de −*c*. (Voir l'article *Nombre entier relatif*).

Prenons maintenant un volume simple, un polyèdre. Ce polyèdre peut être défini par les coordonnées de ses sommets. Par la donnée d'une série de valeurs [ (*x*1, *y*1, *z*1) ; (*x*2, *y*2, *z*2) ; … ; (*xn*, *yn*, *zn*) ], on définit ce volume.

Prenons par exemple les huit points

* (-4, -4, -4)
* (-4, -4, 4)
* (-4, 4, -4)
* (-4, 4, 4)
* (4, -4, -4)
* (4, -4, 4)
* (4, 4, -4)
* (4, 4, 4)

ces huit points définissent les sommets d'un cube dont l'arrête a une longueur 8, et dont le centre est en *O*.

On a ainsi représenté un cube par un ensemble de valeurs. Cet espace est également appelé une matrice tridimensionnelle ou « virtuelle » dans le monde de l'imagerie 3D.

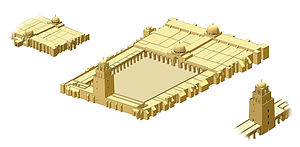
**Techniques**

La synthèse d'image 3D se décompose essentiellement en deux étapes :

* Modéliser ce que l'on veut visualiser ou représenter. Cette étape est appelée **modélisation**
* Effectuer la visualisation de ce que l'on a modélisé. Cette étape est appelée rendu

Dans chacune de ces étapes, un grand nombre de techniques existent. Bien que théoriquement les techniques de modélisation et de rendu soient indépendantes, il va de soi que le rendu doit pouvoir exploiter les données modélisées et bien souvent une technique de modélisation est étroitement associée avec une technique de rendu.

**Modélisation**

[](http://commons.wikimedia.org/wiki/File:3D_computer_modeling_of_the_Great_mosquee_of_Kairouan-blank.jpg)

Modélisation 3D de la Grande Mosquée de Kairouan (en Tunisie).

La modélisation va consister à faire stocker par l'ordinateur un ensemble de données géométriques et de caractéristiques graphiques permettant de représenter ensuite le *modèle*. Ce modèle est habituellement appelé scène 3D, d'après l'anglais *scene* qui signifie « vue », on dira donc « vue 3D ». Voir l'article dédié pour plus d'informations :

Article détaillé : Modélisation 3D.

**Types de modélisation**

* CSG (*Constructive Solid Geometry*, géométrie constructive des solides) : les objets sont définis par des opérations (unions, intersections…) sur des formes géométriques dans l'espace (polyèdres, sphères, cônes, courbes de Bézier, splines, NURBS…). Cette méthode est bien adaptée pour la création *ex nihilo*, puisqu'elle correspond bien à la façon dont l'esprit humain se représente la forme des objets.
* B-Rep (*Boundary Representation*, représentation par frontières) : la surface de chacun des objets est limitée par des formes géométriques 2D (généralement des triangles). C'est le format naturel de sortie des scanners 3D qui mesurent les cotes de points à la surface des objets. C'est aussi la représentation généralement utilisée par les dispositifs d'affichage accélérés pour la 3D.
* Splines, où les surfaces sont décrites par des courbes formant un treillis.
* Surfaces implicites, où les objets à représenter sont décrits par des surfaces joignant des volumes, le plus souvent des sphères.
* Voxels, reprenant en 3D l'analogie des images numériques où l'espace est divisé en intervalles réguliers.

**Logiciels de modélisation**

Le processus de modélisation peut être soit automatique comme pour un scanner 3D, où un programme va créer une représentation informatique à partir d'un objet réel du monde, soit manuel à l'aide d'un logiciel d'édition 3D. Dans ce dernier cas, un infographiste est responsable de la création du modèle. Ces logiciels sont appelés *modeleurs* d'après l'anglais *modeler*.

Article détaillé : Logiciel de modélisation 3D.

Parmi les logiciels les plus connus et les plus répandus se trouvent 3D Studio Max, Maya, XSI, LightWave, Cinema4D, 3D Turbo, Modo et le logiciel open source, Blender.

**Rendu**

Le rendu se décompose à son tour en plusieurs phases :

* Calcul de l'éclairage (on parle d'illumination)
* Projection dans le plan d'observation.
* Dessin à proprement parler avec application éventuelle de textures

Il arrive que certaines de ces phases soient réalisées en même temps (dans les cartes 3D par exemple)

**Moteur de rendu 3D**

* Le rendu est une phase qui consiste à transformer l'espace 3D en une image 2D. Le rendu est généré par un/des programme(s) qu'on appelle Moteur de rendu 3D, intégré(s) ou non au logiciel de modélisation.

**Illumination**

* illumination locale
  + L'algorithme de Gouraud, est une méthode de rendu très simplifiée puisqu'elle consiste à interpoler sur l'image en 2D, ligne par ligne au niveau des pixels, les valeurs d'intensité de l'intérieur d'un polyèdre selon les intensités de ses voisins : c'est un lissage 2D.
  + L'algorithme de Phong est un ombrage 3D qui permet d'avoir un volume et un effet de matière, en fonction de la source de lumière et de propriété de réflexion de la lumière du matériau en chaque point du polyèdre, par interpolation des normales aux sommets.
* illumination globale
  + la radiosité calcule les échanges d'énergie lumineuse entre éléments de surface de la scène en tenant compte de leurs propriétés de réflexion et d'absorption. Cela nécessite de décomposer la scène en éléments finis de surface. Cela permet des éclairages naturellement doux. Cette technique associée au lancer de rayons et à d'autres voisines permettant les inter-réflexions est une technique d'illumination globale.

**Projection**

Les objets ainsi définis par des nombres peuvent ensuite être dessinés sur un écran ; les triplets de valeurs (*x*, *y*, *z*) sont transformés en points du dessin. Ce rendu utilise la notion de perspective :

* perspective conique, avec points de fuite : plus l'objet est loin, plus il est dessiné petit ; ce procédé est parfois qualifié abusivement de « vraie 3D » ou de « vision naturelle » ;
* perspective axonométrique, dont la perspective cavalière et les projections orthogonales (notamment la perspective isométrique, voir aussi *Géométrie descriptive*) : la taille de l'objet ne varie pas avec l'éloignement ; l'éloignement est figuré par un déplacement dans le plan de la figure.

D'autres algorithmes de rendus sont utilisés, tenant notamment compte du fait que les détails et les contrastes s'estompent avec l'éloignement (perspective atmosphérique).

Les techniques de la synthèse d'images 3D ont d'abord distingué les algorithmes de **calcul des faces cachées** qui travaillaient dans l'espace 3D de la scène et ceux du **rendu photoréaliste** qui travaillaient dans l'espace 2D de l'image (pixels). Les algorithmes de rendu actuels réalisent les deux fonctions simultanément.

**Dessin**

À ces types de perspective, on associe un type de dessin :

* Z-buffer (tampon de profondeur des pixels), utilisant l'algorithme dit du peintre (qui peint la scène en partant du fond puis successivement les sujets de plus en plus rapprochés) pour afficher les scènes constituées de polygones ou de structures plus évoluées comme les *nurbs*, et qui gère correctement l'affichage de polygones entremêlés à la précision de l'affichage car le tracé se fait pixel par pixel. C'est une méthode 3D de calcul de faces cachées à l'échelle des pixels car elle conserve la profondeur de chaque pixel.
* Le lancer de rayon (*raytracing*, dont est dérivé le raycasting), simulant le parcours inverse de la lumière de la scène vers l'œil. Cette technique permet de simuler relativement facilement les phénomènes physiques que sont réflexion et réfraction ainsi que les ombres portées mais crée des scènes trop pures et irréalistes. Dans ce cas, la projection est réalisée en même temps que le dessin car tracer un rayon implique de le projeter dans la scène 3D.
* Autres…

Des techniques complémentaires d'application de texture (*texture mapping*) sont utilisées pour le rendu d'effets plus réalistes sans pour autant rendre plus complexes les modèles 3D. Par exemple :

* l'application de textures photoréalistes issues de l'objet réel,
* le placage de relief (*bump mapping*) qui est une perturbation locale des normales à la surface et permet de rendre l'aspect de surfaces granuleuses par exemple.
* les textures d'éclairage

Avec ces techniques apparaissent les problèmes de filtrage, qui est nécessaire pour éliminer les artefacts.

**Variantes**

* La représentation par arêtes, équivalente du dessin au trait, ne représente que les bords des objets de la scène ne nécessitant qu'un algorithme de faces cachées. Cette méthode de calcul dans l'espace 3D est encore utilisée dans de nombreux modeleurs pour avoir un affichage rapide.

**Différence fondamentale avec la 2D**

Au final, on a une image 2D. Cette image aurait très bien pu être dessinée directement en 2D, mais elle a été calculée, générée à partir du modèle 3D.

Le monde de la 3D en images de synthèse est, effectivement, mathématiquement réel. Mais le mécanisme ici est inversé, tel un dessinateur reproduisant sur papier carton une œuvre sculptée sous tous ses angles. L'image en deux dimensions résultant du rendu d'une scène tridimensionnelle n'est autre que le sous-produit de ce monde virtuel « filaire ». Les possibilités de cette technologie sont infinies, tout comme un artiste dessinateur pourrait reproduire la réalité sous une infinité d'angles et de paramètres différents.

Contrairement à une image tridimensionnelle classique, une image tridimensionnelle volumétrique (ou volumique), donne une valeur à tous les points de l'espace (calculs matriciels). Ceux-ci sont tous reliés (ou non) entre eux afin de constituer des polygones que l'on appelle communément des facettes. Ces facettes interconnectées entre elles à leur tour, constituent au final l'objet 3D en question. Par ailleurs, d'autres points peuvent se trouver eux aussi à l'intérieur même de l'objet, ceux-là même auxquels on peut attribuer des fonctions spécifiques supplémentaires, conférant à l'objet une impression de masse en plus du volume (ex. : Objet plein / Objet creux).

**Rendu du relief**

La vision en relief provient du fait que, les deux yeux étant décalés, un objet n'est pas vu au même endroit par les deux yeux (différence de parallaxe). L'ordinateur peut générer deux images différentes, une étant vue par l'œil gauche, l'autre par l'œil droit, et donner ainsi une impression de relief.

Un moyen simple de créer ce « relief artificiel » consiste à générer une seule image, dite anaglyphe, mais contenant les informations pour les deux yeux, en deux couleurs, en général vert et rouge. Le spectateur dispose de lunettes ayant un filtre vert d'un côté (cet œil ne voit que les informations rouges, qui apparaissent en noir, voir *Synthèse soustractive*), et un filtre rouge de l'autre côté (cet œil ne voit que les informations vertes, qui apparaissent également en noir).

On peut aussi travailler avec une seule image initiale, avec une répartition des couleurs judicieuses. Les lunettes décrites ci-dessus exagèrent l'effet visuel sur ces images spéciales, mettant au premier plan le sujet de l'image en vert et l'arrière plan du décor en rouge. On obtient ainsi une illusion de profondeur stéréoscopique, ainsi que de relief.

**Holographie virtuelle**

Rien n'interdit avec les processeurs actuels de calculer en quelques minutes, voire en quelques secondes, des *hologrammes* simples impressionnés sur des films à bonne résolution. Le procédé existe en fait depuis le début des années 1970 (l'*IBM Systems Journal* avait fourni un exemplaire d'un tel hologramme *calculé* dans l'un de ses numéros de cette décennie). Le procédé n'eut pas de suite à l'époque pour des raisons de coût, mais pourrait ressurgir aujourd'hui.

**Contraintes temporelles**

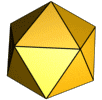
Bien que la technologie soit analogue dans les deux cas (cinéma et jeu vidéo), il existe des différences notables. La contrainte de temps réel inhérente au jeu vidéo n'existe évidemment pas au cinéma. C'est pourquoi les images de synthèse utilisées au cinéma, plus détaillées, sont précalculées une par une, ce qui permet un bien meilleur rendu visuel. Il convient donc de bien faire la différence entre la 3D en temps réel et la 3D précalculée. Le cinéma et le jeu vidéo ne sont pas les seuls contextes d'utilisation de la synthèse d'image 3D. D'autres contextes, notamment celui de la Maquette numérique d'aspect utilisent des solutions hybrides, des compromis entre 3D temps réel et 3D précalculée, qui permettent d'obtenir dans certains cas un rendu visuel proche de la synthèse d'image 3D précalculée tout en bénéficiant des avantages de la 3D temps réel.

**3D temps réel**

La 3D temps réel est effectivement utilisée dans les jeux vidéo, mais a également de nombreuses autres applications : visualisation architecturale, visualisation médicale, simulations diverses, économiseurs d'écrans… Le challenge technique inhérent à cette sorte de 3D est d'obtenir la meilleure qualité d'image possible tout en conservant une animation fluide, ce qui demande d'optimiser au mieux les calculs d'affichage. Au début, tous les calculs revenaient au processeur central (CPU) des ordinateurs, mais la puissance toujours plus grande demandée pour améliorer la qualité des images poussa des constructeurs à commercialiser des cartes PCI (Peripheral Component Interconnect) spécialisées dans la 3D. La première disponible sur PC (Personal Computer) fut la légendaire Voodoo de la société 3DFX. Aujourd'hui, les cartes graphiques intègrent pour la très grande majorité des fonctions d'accélération 3D.

L'utilisation de processeurs dédiés au calcul de la 3D créa la nécessité de définir des API (application programming interface) standard, permettant aux développeurs de facilement accéder aux fonctions accélérées et aux constructeurs de les inclure dans les processeurs. Les deux API 3D les plus répandues sont OpenGL et Direct3D (composant DirectX), concurrent plus récent développé par Microsoft.

**3D précalculée**

[](http://commons.wikimedia.org/wiki/File:Ikosaeder-Animation.gif)

3D précalculée

La 3D précalculée est utilisée dans la création d'images, d'effets spéciaux et de films d'animation. Son avantage principal est qu'elle permet d'obtenir une très grande qualité d'image et un réalisme poussé. L'image est qualifiée de « photoréaliste » lorsque son niveau de détail est très important. On peut alors la confondre avec une photographie. En 2001, *Final Fantasy : les Créatures de l’esprit*, est le premier long-métrage ayant pour ambition le photoréalisme.

La 3D pré-calculée ne permet aucune interactivité. Le spectateur ne peut agir sur la scène 3D. Elle est notamment utilisée au cinéma ou en télévision.

Une préoccupation pour les temps de calcul existe, mais à une échelle totalement différente de celle de la 3D temps réel. En effet, dans cette dernière, une animation fluide demande qu'il soit possible de calculer plus de vingt images par seconde, alors qu'en 3D précalculée, le calcul d'une image peut prendre des heures, voire des jours. Une fois toutes les images calculées, elles sont projetées à la fréquence voulue (24 images par seconde pour du film 35 mm par ex.). Elles sont souvent générées sur ce que l'on appelle une ferme de calcul. C'est un réseau important d'ordinateurs reliés les uns aux autres pour multiplier la puissance de calcul. Comme chaque seconde d'animation requiert près de 30 images, cela signifie que pour 15 minutes d'animations le réseau devra générer plus de 27 000 images.

Les logiciels calculant ces images, les « moteurs de rendu » (*renderers* en anglais) sont nombreux. Au fil des années, les logiciels grand public se sont considérablement améliorés pour affiner le grain et aussi pour transmettre les effets réels sur l'environnement imaginé par les équipes de concepteurs. Par contre, des firmes spécialisées avec des artistes aux connaissances pointues en modélisation, en éclairage, et en animation sont toujours nécessaires. Les autres logiciels les plus connus sont RenderMan, Mental Ray, finalRender, Brazil r/s, V-Ray.

Une technique populaire de calcul 3D est le « lancer de rayon » (*raytracing*). Dans cette technique, un rayon lumineux « part » de chaque point de l'écran, et « rencontre » les objets dispersés dans la scène 3D et se « reflète » sur eux.

La 3D pré-calculée connaît un succès fulgurant auprès du grand public, notamment grâce à sa diffusion au cinéma. Le studio d'animation Pixar est voué uniquement à la création de films en images de synthèse depuis les années 1980. Il est reconnu comme précurseur en la matière. Il développe notamment RenderMan, un moteur de rendu considéré comme l'un des plus performants aujourd'hui, servant au rendu de toutes les productions du studio ainsi qu'à celui de la vaste majorité des images de synthèse de longs métrages (voir films en images de synthèse et Synthèse d'image).

Les domaines de l'architecture, de l'aménagement intérieur ou encore de la publicité utilisent de plus en plus ces méthodes pour faciliter la vente de futurs projets.

**3D hybride**

[](http://commons.wikimedia.org/wiki/File:Roller01.jpg)

Image rendue par Patchwork3D en 3D temps réel avec éclairage pré-calculé

La 3D hybride vise à améliorer le photo-réalisme de la 3D temps réel en pré-calculant (et donc en figeant) certains paramètres dont le calcul est coûteux mais dont la qualité est indispensable au réalisme de l'image.

L'éclairage fait partie de ces paramètres coûteux. En 3D temps réel, il est sacrifié : pas ou peu d'ombres portées, pas d'illumination globale. Or, un bon éclairage apporte énormément au réalisme de l'image et c'est souvent ce qui fait la grande différence entre la 3D temps réel et la 3D pré-calculée. Figer ce paramètre peut être un problème, il ne peut plus être changé en temps réel, l'éclairage ne peut plus être complètement dynamique. Mais il y a de nombreux contextes d'utilisation de la 3D temps réel dans lesquels il n'est pas indispensable d'avoir un éclairage complètement dynamique et ou on peut donc se permettre de figer cet éclairage.

On calcule alors des textures d'éclairage (lightmaps) qui encodent l'éclairage. Ce calcul est relativement long (précalculé, il n'est pas fait en temps réel) mais, une fois effectué, l'information d'éclairage est disponible instantanément. La qualité du rendu peut alors approcher ce qu'on obtient en 3D précalculée tout en apportant de l'interactivité dans les réglages de tous les autres paramètres.

[](http://commons.wikimedia.org/wiki/File:Bump.gif)

Changement d'un matériau en temps réel

**Salons**

* Le Siggraph est le rendez-vous annuel des laboratoires de recherche sur la synthèse d'image.
* Chaque année, le festival des **images de synthèse** Imagina a lieu à Monaco.