

Réalité Enrichie par Synthèse

Jancène Pierre, Fabrice Neyret, Xavier Provot, Jean-Philippe Tarel, Jean-Marc Vezien, Christophe Meilhac, Anne Verroust

▶ To cite this version:

Jancène Pierre, Fabrice Neyret, Xavier Provot, Jean-Philippe Tarel, Jean-Marc Vezien, et al.. Réalité Enrichie par Synthèse. 10ème congrès AFCET, Reconnaissance des Formes et Intelligence Artificielle, Jan 1996, Rennes, France. pp.923-931. inria-00588889v1

HAL Id: inria-00588889 https://inria.hal.science/inria-00588889v1

Submitted on 26 Apr 2011 (v1), last revised 13 Jul 2012 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RAitAiAiA

Pierre JancŁne, Christophe Meilhac, Fabrice Neyret, Xavier Provot, Jean-Philippe Tarel, Jean-Marc VØen, Anne Verroust

INRIA

Domaine de Voluceau, ROCQUENCOURT 78153 Le Chesnay cedex FRANCE Contact e-mail: Christophe.Meilhac@inria.fr http://www-rocq.inria.fr/syntim/analyse/video-fra.html

R@mØ

Nous proposons une modode pour automatiser l'insertion comon d'objets 3D de synthLse dans des so quences d'images rolles. Celle-ci fait appel conjointemendes techniques d'analyse et de synthLse d'images pour calculer les images modifion.

Mots-clØ

analyse d'images, synthLse d'images, raitØirtuelle.

1 Introduction

1.1 Etat de l'Art

La rolit@ugment@prend de l'ampleur dans diffonts domaines comme les processus de fabrication pour l'industrie [1, 2], l'architecture d'int@eur [3, 4] et le domaine molecal [5]. Dans la plupart de ces applications, le problLme principal est d'assurer la provision de la superposition des images calcul@s sur les images rolles.

En effet quand on insŁre des objets 3D de synthŁse dans une source d'images pour des applications de vid \emptyset il faut s'assurer de la cohorne de l'illumination, des ombres et des occultations des objets virtuels par les objets r \emptyset s. Il existe plusieurs approches en r \emptyset onse à ce problŁme :

- Les applications actuelles d'effets sp@aux pour la production vid@superposent simplement l'image de synthLse devant l'image r@le (incrustation ou *alpha keying*). Si l'image de synthLse doit de partiellement occult@ un masque 2D doit alors de fourni, ce qui interdit l'animation ou contraintimaterpoler l'@olution du masque au cours du temps.
- La Synthetic TV [6], garantit la cohéance gốmØ trique entre images rélles et objets de synthLse en synchronisant les mouvements de la 'cam@a' de synthLse, correctement modØs@ avec les mouvements de la cam@a rélle.
- La *Computer Augmented Reality* [7, 8], s'attache au calcul de l'illumination globale dans une scŁne rØle enrichie d'objets de synthŁse.

Ces méthodes nécessitent une forte interaction avec l'utilisateur lors de l'opfation de mixage. Notre but est d'automatiser partiellement ce traitement.

Abstract

In this paper, a technical solution is presented to automate the mixing of real and synthetic objects in a same animated video sequence. We aim at achieving a close binding between 3D-based analysis and synthesis techniques to compute the interaction between a real scene captured in a sequence of calibrated images, and a computer-generated environment.

Keywords:

Image analysis, Image synthesis, virtual reality.

1.2 Notre Approche

La Réditéonrichie par SynthLse (RES) a pour but l'insertion d'objets 3D de synthLse dans une séquence 2D d'images rédies, en assurant la cohéonce 3D pour les occultations, les ombres, les collisions et les contacts. L'idéoconsistercicontiné et localiser des objets 3D dans les images 2D, afin de construire des sortes de masques 3D.

Notre approche associe des techniques d'analyse et de synth Les d'images via la constitution de mod Les 3D synth Sables compatibles, autorisant un mixage 3Dcoh Sables de synth Les dans les images 2D.

Les applications de cette technique peuvent se trouver en visualisation (par exemple en imagerie m@icale), et pour la production vidØ (effets sp@iaux pour les films, publicitØ@ide d'impact...).

Un modLle synthØsable dØit une scLne composØ d'objets, de sources de lumiLres et d'une camØa. Les objets ont une forme, une position, une orientation, des propriØde surface (ex : couleurs ambiante, diffuse et sp@ulaire) et Øentuellement une Øblution temporelle (mouvement rigide ou dØrmation). Un modeleur est un outil interactif permettanuà utilisateur de dØire un modLle synthØsable 3D.

Les traitements effectu $\not\!\!\! \mbox{0}$ pour la RES sont effectu $\not\!\!\! \mbox{0}$ en diff $\not\!\!\! \mbox{0}$

- La source rolle est tout d'abord numøsøet analysø: les modLles des objets 'reconnus' dans la scLne rolle sont recalø(cf. partie 2), de maniLreroronstituer une scLne synthøque les remettant aux places et orientations identiquescèlles de la scLne originale. Les caractøstiques de la camø sont øalement estimøs.
- L'utilisateur sp@fie sa scŁne anim@de synthŁse à l'aide de notre modeleur ACTION3D (cf. partie 3), en partant de la scŁne recal@construite par l'analyse.
- Pour chaque image de la some, le processeur de production d'images lit les donnes gomon d'action d'images lit les donnes gomon d'action d'action

Ceci conduit aux sources d'images protentos dans la partie 5, montrant l'insertion d'objets 3D de synth£se dans des images ou sources rolles avec maintien de la cohorne 3D pour les occultations, les ombres et les collisions.

2 Analyse d'image

2.1 Introduction

L'une des sp@ficitØde l'application RØitØnrichie par la SynthŁse (RES), par rapport aux applications plus classiques d'analyse comme celles do charobotique, est la næssitø'obtenir une description de certains objets observøplus complŁte que pour les applications d'analyse classiques, tant dans son aspect $g \phi$ møique que photomøique. En particulier, la connaissance gomoique noessaire pour pouvoir ajouter des objets de synth¿se en interaction 3D raiste avec la scLne døasse de loin les donnøs que peuvent extraire les modes de reconstruction connues en vision par ordinateul'àeure actuelle. Nous avons donc choisi une m@iode d'analyse par modŁle, qui suppose d'introduire des connaissances gomoriques donnos a priori sous la forme d'un modèle des objets 3D proents dans la scŁne. Ceci permet de restituer les parties importantes de la scŁne de fac robuste, pàrtir d'informations dans les images.



FIG. 1 - Qu'est ce qu'une scEne?

Dans notre cadre, une description du monde contient un observateur, des sources lumineuses et des objets (figure 1). Le but que se fixe l'analyse est de fournir les



FIG. 2 - Le portique du laboratoire.

informations næssaires (gønøie et photomøie) sur ces trois types d'entitøpræntes dans la scŁne observø La description de l'observateur est rølisøen ligne par calibration (section 2.3). Les modŁles des objets sont positionnøpar recalage sur des donnøi 3D extraites des images (section 2.4). Enfin, on identifie les sources de lumiŁres sous la forme de conditions d'ølairage øluvalentes (section 2.5).

La description, ou mod Łle de la sc Łne, est constitu $\not\!\! {\mathfrak Q}:$

- 1. d'un aspect g@m@ique : localisation des objets prØ sents dans l'image dont on a un modŁle, modØsation et mouvement de la cam@i,
- 2. d'un aspect photom@ique : localisation des sources de lumiŁre @uivalentes et estimation de la r@ectance des facettes.

2.2 Contraintes sur la scŁne

Nous réapitulons dans cette partie les contraintes dues au matédel (cf. fig 2) et les hypothLses qui simplifient le traitement. Ces hypothLses limitent la géfédlitØ du systLme, la relaxation de ces contraintes constitue une part de notre travail actuel. Cependant, l'objectif de cet article est avant tout de dédontrer la faisabilitØ de la RES, c'estdare la possibilit@l'ajouter des objets de synthLse dans une scLne r@le en interaction 3D rédiste, de maniLrad@asser les limites des systLmes d'incrustation actuels.

2.2.1 Contraintes d'acquisition

Le laboratoire dispose d'une camt couleur Sony XC007 avec un objectif zoom Canon J15x9.5B. La camt est plac sur un portique qui permet le pilotage du zoom, le r gage du diaphragme de l'objectif et son d camt selon un axe de translation et deux axes de rotation (figure 2).

Un mouvement de camé calibr Puisque le systLme d'acquisition est pilotable avec une bonne présion, nous supposons que la position de la camé en rotation et translation ainsi que son zoom et sa mise au point sont connues. Ces informations sont accessibles dLs que le portique est calibr étimé atiquement [9]. Dans l'ét actuel de nos déeloppements, seuls les dé placements en translation de la camé sont calibré

Des scLnes statiques. Les traitements sur les images d \emptyset rits par la suite (partie 2.4 et 2.5) doivent **Ge** assez pr \emptyset s pour permettre une incrustation de bonne qualit \emptyset

sinon on risque d'observer des effets de bavures entre les images des objets rØs et de synthLse. Nous supposons donc que l'image observØest propre, et en particulier qu'elle ne contient pas trop de bruit, ce qui nØessite une moyenne sur plusieurs images d'un mØfe point de vue. Puisque nous ne nous intØssons qudès scLnes statiques, il suffit de faire l'acquisitionlàcadence de B`àmages par seconde.

Pas d'analyse en temps rØ. Comme indiquén introduction, la RES se døule dans un contexte 'diff@Ø et non 'temps rØ', ce qui permet de traiter des sØ quences prØxistantes calibr@s, et de proc@erdas op@ations d'analyse complŁtes donc conteuses. Dans l'Øat actuel de notre systLme, il faut compter une durØ d'une journØpour analyser avec succLs une paire stØ rØscopique extraite d'une sØnence de deux secondes.

2.2.2 HypothLses simplificatrices

Comme il n'est pas possible d'obtenir toute l'information g@m@ique utildaesynthLse d'images sur un objetraeins de disposer de tous les points de vue, seuls quelques indices 3D robustes sont extraits des images pour permettre le recalage avec les modLles complets des objets de la scLne.

ModŁle des objets. On suppose lors du processus d'analyse des images qu'un modŁle des objets les plus importants dans la scŁne est disponible.

Le modLle g@m@ique d'un objet est d@ni par ses surfaces les plus caract@stiques. La description utilis@ est une repr@ntation par bords 3D, ou les surfaces sont d@nies comme des polyLdres, eux-m@fes repr@nt@ par un ensemble de polygones orient@ ou facettes. Les parties courbes d'un objet sont approximables par ce type de repr@ntation. Ainsi, certaines parties des objets manipul@peuvent @e courbes, mais elles ne seraient pas utilis@ pour l'analyse st@scopique. Cette reprØ sentation est complLte du point de vue g@m@ique, puisque le point de vue et les conditions d'@airage n'interviennent pas dans la caract@sation de l'objet.

Des facettes planes. Nous supposerons que les objets observødont on døre obtenir une localisation posskdent dans l'image suffisamment de surfaces planes visiblesphotomøre uniforme. Un minimum de trois facettes visibles et non occultør est nørssaire pour chaque objet afin d'obtenir un recalage robuste et prørs de son modkle gømørique. Cela induit les hypothkses suivantes :

- les conditions d'illumination doivent de suffisamment bonnes : Ølairage lointain, diffus et important ;
- 2. les mat@aux doivent file de nature peu sp@ulaire de maniŁrer@aure les reflets de lumiŁre.

Ces hypothLses permettent de røiser une analyse en rø gion et d'obtenir de maniLre directe, par reconstruction, des facettes 3D (voir section 2.4). De plus, la face est la primitiveagomøie la plus simple qui peut contenir une information photomøique de nature texturelle. On suppose donc que la scŁne observØest localement plane : cette hypothŁse est toujours vØfiØsi l'on observe des surfaces suffisamment petites, mais raisonnable uniquement pour des scŁnes de type 'intØeur', comme des scŁnes de bureau ou des scŁnes architecturales.

2.3 Analyse du point de vue

Dans cette partie, nous explicitons la formation d'une image, et comment on acquiert les connaissances qui concernent l'observateur (position, mouvement, caractØstiques propres).

L'observateur est digit par un ensemble de paramLtres. Ceux-ci sont appel Ø*point de vue*. Celui-ci est modifiable par les paramLtres de conterôlu portique : commande de translation, rotation, zoom. La calibration, c'estdare la connaissance de la transformation que subit un point de vue par les paramLtres de conterô est rídis Ø ar processus hors-ligne grâne mire de calibration (objet de gØmØ ie parfaitement connue). Ainsi dLs l'acquisition, le point de vue de l'observateur est connu.

2.3.1 Difficition d'un point de vue

Le point de vue est d**Ø**rmin**Ø**ar :

- une origine, appelØcentre de la camØ,
- une direction de visØ dont l'axe passe par l'origine,
- un rectangle image, le plan image est orthogonalà direction de vis€
- un twist, qui fixe l'angle de l'un dest **Sti**e l'image avec l'horizontale,
- une focale et
- une distance de mise au point, les objets qui ne sont passette distance sont flous.

Un certain nombre d'actions sur la camé peuvent modifier les paramŁtres du point de vue (le focus ou mise au point, le zoom, le diaphragme et les délacement du robot en rotation et translation).

Nous utiliserons une version simplifié de la modéstion d'une camé a, le modèle stéopé cf. figure 3). L'image d'un point dans l'espace euclidien 3D est l'intersection du plan image et de la droite passant par le point objet et le centre. L'objectif de la calibration est d'extraire les paramètres de l'observateur et d'ére en mesure de suivre les transformations de ces paramètres sous l'effet des commandes du portique.

2.3.2 Calibration de l'objectif de la camé

La calibration d'une camé est une métodologie permettant d'extraire les param£tres du mod£le. Elle est rédiséden généal genêa mire de calibration (objet de généaie parfaitement connue). On rep£re la position d'un certain nombre de points caractédstiques sur la mire par rapportune rep£re ligealle-ci, et on extrait dans l'image de calibration le point image correspondant hacun d'entre eux [10].

Une cam@ possLde deux commandes pour am@orer la qualit@de l'image, qui ont des cons@uences sur la projection perspective : le zoom et le focus. Chacune



FIG. 3 - ModŁle støpp@'une cam@.

d'entre elles modifie le point de vue : le centre se translate sur l'axe de visØet la focale augmente ou diminue.

2.3.3 Calibration du portique

Pour Øddier le mouvement de la camØd, on peut distinguer trois types de repŁres. Le premier est liØu monde donc immobile lors d'un dØdacement de camØd, le second est le repŁre camØd dans lequel la projection centrale sur le plan image s'exprime naturellement, le dernier est le repŁre pince dans lequel les commandes de dØdacement sont connues.

La calibration cam@-pince est une m@iodologie qui doit fournir la matrice de passage entre le repŁre pince et le repŁre cam@i. L'article [9] pr@ente une m@iodologie de calibration.

2.3.4 R@apitulatif

Nous venons de dérire notre défarche pour obtenir un systLme entiLrement calibré dLs l'acquisition. Nous avons défompos d'analyse du point de vue en deux parties : d'une part l'extraction des param Ltres life à camé a elle-mére (les commandes robotiques susceptibles de les modifier sont le zoom et le focus) et d'autre part l'extraction des param Ltres défivant le mouvement de la camé en translation et rotation. Dans les deux cas notre calibration repose sur des donnés images et sur l'observation d'une image de mire.

On peut amorer notre systLme de calibration en rehânt un certain nombre de contraintes tel que la restitution du mouvement par des calculs sur l'image, ou encore se passer totalement de mire de calibration, les paramLtres de modosation cont entiLrement (ou partiellement) calculoritation des images, a posteriori [11].

2.4 Analyse de la position des objets

Une fois les caract destiques de l'observateur obtenues, il est possible d'extraire des informations 3D sur les objets de la scŁne. Comme nous connaissons par hypothŁse un modŁle des principaux objets qui constituent la scŁne, le processus d'analyse des paires d'images consiste àtraire suffisamment de primitives 3D pour localiser robustement ces objets dans la scŁne. Ainsi, dans notre processus chaque modŁle des objets importants est recal@r@s@ent par rapport aux facettes 3D reconstruites. Seuls les objets qui peuvent de repr@nt@par un ensemble de contours rigides et qui possŁdent suffisamment de surfaces homogŁnes planes sont actuellement recalables par notre systŁme.

2.4.1 Extraction de caractØstiques 3-D

Dans le processus d'analyse de la source, nous choisissons des paires d'images ou l'objet qui nous intoris de set observoir deux points de vue assez diffonts pour permettre une reconstruction de qualitor L'objet doit tre bien visible dans l'image pour pouvoir effectuer le traitement dans de bonnes conditions. La position des objets dans l'espace 3D n'est alors calculor que sur ces deux images extraites de la source.

Notre m@tode repose donc sur une analyse st@00 scopique,p\u00e0rtir d'une paire d'images gauche-droite, dont les @tpes sont (figure 4):

- **Segmentation :** chaque image est døbupøen rølons, indøendamment de l'autre image,
- Appariement 2D: Les rgions gauches identifigit prØ cgemment sont si possible mises en correspondance avec les rgions de l'image droite,
- Reconstruction 3D: les rgions, de chaque appariement issues d'une marie surface plane, sont localisgi dans l'espace 3D euclidien.

Ces @ pes nous fournissent un ensemble de primitives 3D qui nous permettra par la suite de faire du recalage. Le but n'est donc pas d'obtenir une reconstruction exhaustive mais plut@ avoir suffisamment d'@ ants reconstruits pour que le recalage 3D soit possible avec une bonne pr@ sion.

Nous allons dØrire en dØil les diffØntes Øpes.

a) **Correction des images Gréa**ux informations fournies par la calibration, les images du couple statiscopique sont corrigés de leurs distorsions pour s'approcher d'un mod Lle de projection statiop@arfait [10]. Les distorsions photom@iques doivent Gré aussi corrig@s pour ne pas biaiser les algorithmes qui s'appuient sur les intensit@dans l'image [12].

b) Segmentation en rgions : primitive 2-D Les paires stocopiques calibros sont segmentos en rgions, afin d'extraire les principales zones homogLnes au sens d'un certain critLre, qui seront reconstruites sous forme de faces. Ces facettes servent d'ancrage au recalage des modLles des objets les plus importants de la scLne. Le principe de la segmentation est de regrouper progressivement les pixels de l'image avec leurs voisins en fonction de critLres d'homogatoccei peut Ge rgo lis@ar croissance de rgions [13, 14], ou par division roursive de l'image pilot@par la mise en correspondance [15, 16], par classification floue [17] ou par une minimisation d'orgie [18].

L'id d serait un processus de segmentation aussi peu param d que possible (par exemple, ne d dendant



FIG. 4 - Le processus d'analyse statiscopique.

que d'un paramLtre d'Øhelle), afin d'automatiser la procØure. De ce point de vue, la derniLre mØlode conduit dès rØultats facilement contatôles.

c) Mise en correspondance 2D Comme pour des primitives plus classiques, chaque rgion, une fois segmentgi doit Gre apparigiavec son correspondant dans l'autre image de la paire. Un des avantages de l'approche par rgion tientarquantitgi information disponible pour opger les bons appariements : des critŁres ggingiques (forme, contrainte gapolaire), mais aussi photomgiques (l'aspect de la projection d'une rgion change peu entre deux points de vue stagi peuvent Gre utilisgi

Pratiquement, la médode mise au point cherche, pour chaque région de l'image gauche, un ensemble de régions candidates dans l'image droite, suivant un critŁrechâ de ressemblance photomédique (variance et moyenne similaire). On trie ensuite les candidats par un ensemble de contraintes : taille semblable, contrainte édipolaire sur le centre de gravit@13]. Cette technique conduit assez facilement aux bons appariements pour les principales régions de l'image, médie en préence de quelques erreurs de segmentation locales.

Exacore, comme le choix des critŁres et des seuils attachøpeut poser problŁme, nous avons døni une

strat@ie syst@iatique qui conduitd`a bons r@iltats pour une gamme de scLnes d'int@ieur. Cependant, il serait possible de raffiner le processus selon le contenu d@ict@ians les images.

d) **Reconstruction 3D** Nous disposons de rgions apparige, pour lesquelles une triangulation directe est impossible. Une mise en correspondance explicite des points extraits des rgions aurait pu rgoudre cette difficult@Des points particuliers des contours, comme par exemple ceux de forte courbure, conviennent bien. Mais cette approche souffre souvent des erreurs de segmentation intervenues lors des @apes de traitement basniveau.

C'est pourquoi, nous avons dœlopp@ne m@node de reconstruction de faces 3D, qui est globale sur les rØ gions appari@s. Les facettes sont reconstruites comme des morceaux de surfaces planes par le biais de leur @uation. Quant aux frontiŁres des r@ions 2D, elles servent uniquementd@ouper le contour 3D dans le plan calcul@19, 20].

2.4.2 Recalage 3D

A cette dape du processus d'analyse, nous avons à notre disposition les modLles 3D d'objets susceptibles d'de pronts dans la scLne. Gabilité de reconstruction 3D doite profilemment, nous disposons de plus d'un ensemble de facettes planes profintes et bien localisos sur les objets. Retrouver la position de ces objets se rouse dans l'un et l'autre de ces deux ensembles de faces. Deux types de monodes ont dois veloppos et sont utilisos de maniLre complonentaire :

Le premier algorithme est døvøe la mølode de recalage ICP [21]. Il s'agit d'un algorithme itøtif qui calculæñaque pas une mise en correspondance de chaque point de donnøavec le point le plus proche du modŁlepòsitionner. Cette correspondance permet le calcul d'un dølacement optimal entre le repŁre de reconstruction et le repŁre propre a l'objet au sens des moindres carrø L'algorithme itŁre les deux øpes appariement/recalage jusqucà que l'erreur røduelle soit stable. A partir d'une position initiale, une telle procøure converge systølatiquement vers le plus proche minimum local.

Les mises en correspondance sont analys statistiquement daque @pe de l'algorithme pour ne retenir que les plus robustes. Ceci a l'avantage de permettre le recalage d'un mod Lle d'objet isol@ans une scLne complexe mone si la reconstruction est partielle, puisque les fausses mises en correspondance sont alors en grande partie ignor [19, 22].

• Les méthodes de classification floue, de par leur robustesse, fournissent un outil intérssant pour la reconnaissance 3D. Une méthode basé sur l'algorithme de classification avec du bruit a édéceloppé qui permet de rédiser un recalage grossier rapidement sur les faces. Le recalage est ensuite affinérar l'algorithme dérit au point préfient. Cette méthode a l'avantage de géer en parall&le plusieurs hypoth&ses de recalage intérssantes en imposant une contrainte globale de bon recouvrement de l'ensemble des dØ placements possibles [23].

2.5 Analyse des sources lumineuses

L'illumination de la scEne est mod ØsØ par un flux lumineux parallŁle qui provient d'une source situØlänfini. Elle est ainsi uniquement dØrminØ par sa direction d'illumination et son intensitØumineuse. Le calcul de l'Øairement sur chaque facette est donc particuli¿rement simple, ce qui permet, en minimisant l'Øart entre la distribution d'intensitØprØue et les intensitØeffectivement observØs sur l'intØeur des rélons, d'obtenir une estimation de la source éuivalente et de réflectance des surfaces. Le métre type d'approche a *Mappliqué* modèles de sources plus complexes comme une source ponctuelle uniforme Quivalente [24]. Evidemment, l'obtention d'une source lumineuse quivalente ne permet pas de faire des estimations ayant une grande valeur du point de vue strictement physique, mais elle fourni des roultats suffisants dans le cadre de la RES.

2.6 Reapitulatif

Nous avons proentoin ensemble d'outils qui permettent d'analyser une souence d'images rolles dans le but de fournir un modŁle g@m@ique et photomØ trique de la scŁne suffisamment riche pour Ge utilisØ dans le cadre de la RES. Nous avons défit les possibilitødu systŁme actuel et les contraintes que cela implique sur la nature des objets observøet du systŁme d'acquisition des images. En particulier, les objets doivent poss@er des facettes planes uniformes et ne pas se dølacer dans la sølence. De plus, la camøl et son mouvement doivent de calibrø Enfin, les sources lumineuses doivent fe ØignØdes objets ØairØC'est dans ce contexte, que les processus d'analyse du point de vue, d'extraction de facettes 3D, de positionnement des modŁles d'objets, et d'estimation d'Øairage Øuivalent, protemment dorits, peuvent fonctionner.

3 SynthLse d'images

3.1 Introduction

Comme indiqu@r@@emment, le modŁle synth@sable d'une scŁne comprend des objets, des sources de lumiŁre et une cam@a (point de vue).

La problédatique de l'analyse d'images est de reconstituer un mod Le aussi complet que possiblepàrtir d'images 2D. Les problédatiques de la synth Les d'images sont l'inverse de celles de l'analyse d'images. Plus prédsédent, la synth Les d'images s'intédese au probl Les direct de la simulation des divers phéom Les qui conduisent de la description d'une sc Les production d'une image [25]:

modifisation : Une premiŁre probl@atique concerne la description g@m@ique des objets, puis la structuration d'ensembles d'objets. Les modalit@de description sont trŁs vari@s : la modifisation peut ffe interactive, constructive, analytique, sous contraintes. Il faut @dier non seulement les diverses reprontations, mais aussi les procodures conduisantcès mod disations : dormations globales, 'sculpture' par modifications successives, grammaires.

De more, les types de structures liant les objets sont multiples: hi@archies, articulations, squelettes. Il existe @alement nombre de modLles d@i@es types d'objets particuliers (systLmes de particules, liquides, fum@s et nuages, objets complexes, fractales). Le logiciel permettant de constituer cette description est un modeleur g@m@ique.

animation: Une seconde problétatique est libé volution temporelle de ces objets [26], qui comprend la cinétatique, la simulation physique, l'animation des déformations, le conterêdu mouvement, la simulation comportementale. Le logiciel permettant de spétifier ces animations est un modeleur d'animation.

rendu : Une troisiŁme probl*A* atique s'attachel**à** simulation des interactions lumineuses et au calcul de l'image proprement dite, et se s*A* are en deux branches, l'illumination globale et l'illumination locale¹. L'illumination globale prend en charge la r*A* ution de l'*A* uilibre dans les *A* anges d'*A* ergie lumineuse, tandis que l'illumination locale s'occupe de l'interaction de la lumiŁre et de la matiŁre (mod*A* sation de la r*A* ectance, des textures...). Le logiciel effectuant le calcul d'une image est un programme de rendu (ou render). Il est g*A* alement con*A* are de reconstituer l'*A* te la scŁne c**b** aque pas de temps puis de post-traiter et de stocker les images apr*L* s calcul.

3.2 SynthLse d'images pour la RES

3.2.1 construction de modŁles pour l'analyse

Notre modeleur ACTION3D est tout d'abord utilis@pour construire interactivement les modLles g@ m@iques des objets r@s avec lesquels on souhaite interagir. On utilise plusieurs images de l'objet'd@ calquer' que l'on place dans deux vues sur lesquelles on reprojette le modLle en cours de construction. Un systLme permet de limiter le placement des nouveaux points successifs dans les plans des faces d@quises [27].

3.2.2 construction de modŁles pour la synthŁse

Le modeleur est par la suite utilis@lassiquement pour modØser et animer les objets de synthLse rigides ou dØrmØg@m@iquement. Les objets rØs par les lois de la dynamique sont construits par des modules dØfØdirectement lors de l'exØntion du rendu. Le modeleur ACTION3D permet de :

- construire interactivement des objets poly@riques et des surfaces de forme libre,
- dørmer et animer ces objets par des mødodes de dørmation de l'espace 3D et par des mødodes de mødorphose 3D,

^{1.} Sans compter la mod \mathcal{Q} sation des sources de lumi Ł
res et des cam \mathcal{Q} s.

• construire interactivement les textures, et spøffier leur plaquage sur les surfaces.

3.2.3 modŁles dynamiques pour la synthŁse

Dans l'exemple prodet de partie 5, on utilise un module de simulation dynamique de tissus, constitu d'un rocau masse-ressort qui rocation de force et aux collisions avec d'autres objets [28]. A chaque pas de temps, la nouvelle position des masses est calculor fonction des forces internes (mod de de par les ressorts et la dissipation) et des forces externes (vent, collisions...). Il n'y a aucune difforne de traitement selon que le tissu rencontre un objet virtuel ou le mod Lle d'un objet reconnu.

3.2.4 calcul du rendu

Comme indiqu@lus haut et plus amplement d@it dans la partie 4, quatre images de synthŁses partielles sont calcul@s, puis mix@ avec l'image r@le.

Pour le calcul du rendu, nous utilisons le logiciel de rendu du domaine public RAYSHADE, qui opŁre par lancer de rayons [25] : l'øergie ræçpar chaque pixel du plan image est ØaluØen lamç quelques rayons via ce pixeltràvers la scLne. Ces rayons 'remontent' les trajectoires des photons en partant de l'observateur. L'illumination locale est ØaluØau point d'impact de chaque rayon sur un objet de la scŁne, en fonction des propricode surface de l'objet et de la lumiLre parvenant en ce point, selon un modŁle de røcetance qui indique quelle proportion de l'élergie est élise vers l'observateur. L'ørgie incidente est elle-more øalugen lant un rayon vers les sources de lumi
Lre, afin de tester les occultations (gélant les ombres). Un rayon est galement lancgalans la direction de refection donnØpar les lois de Descartes, afin de collecter l'øergie d'un øentuel reflet.

4 Mixage (compositing)

En incrustant (figure 5) l'image de synthŁse sur l'image r@le, l'occultation d'objets r@ls par des objets virtuels @it d@fective dLs les premiLres approches pratiqu@s en vid@ Un masque associ@image de synthLse indique les pixels occup@(avec une r@lution subpixel), contarôt le mixage des images. Celui-ci peut fre obtenu en effectuant un calcul de rendu aprLs avoir peint les objets en blanc uniforme (i.e. isotrope et non lambertien).

En fusionnant les modèles reconnus avec les modèles de synthèse, l'occultation des objets virtuels par les objets réas est automatique avec notre représentation. Il faut cependant en tenir compte pour la gélé ration du masque : les objets reconnus ne doivent pas appatré dans l'image de synthèse ni dans le masque, dans la mesure ou ils sont déprésenté dans l'image réale, mais ils peuvent némeoins cacher en partie des objets de synthèse. Ceci est obtenu en peignant les objets reconnus en noir et les objets virtuels en blanc uniforme. Les objets reconnus sont alordar fois invisibles et occultants, et peuvent gentuellement de réachissants.



FIG. 5 - Noter la trace des objets røs reconnus (la ho), invisib les mais ombrants et occultants.

Les ombres (figure 6) des objets røls sur les objets virtuels apparaissent automatiquement avec notre repræntation unifiø Les ombres des objets virtuels sur les objets røls sont obtenues en calculant un coefficient d'attønationnaultipliel'amage originale. Celui-ci est øalu@en effectuant un rendu des objets reconnus en blanc lambertien, avec et sans les autres objets, afin d'øaluer l'effet d'ombrage d@ aux objets virtuels. Le facteur d'attønation est le ratio de l'image avec les ombres et de l'image de røøence contenant seulement les objets reconnus. (Cette valeur obtenue pour tout point de l'image est øalement calculøavec une røblution subpixel.)



FIG. 6 - objets reconnus, seuls ou avec les objets de synth \mathcal{L} se ombrants. L'assombrissement est donn \mathcal{G} ar le ratio des images.

Le mixage (figure 5) est alors d α ni par : image enrichie = image r α le × att α uation × (1-masque) + image de synthLse.²

5 Exemples

Les figures 7 et 8 sont extraites de sources d'images enrichies.

On peut remarquer dans la figure 7 les diverses formes d'interaction 3D entre objets virtuels et objets r \mathfrak{A} s (ici la **tie**) : appui de la nappe sur la **tie**) ombres de la nappe sur la **tie**) et de la **tie**) sur la balle verte, occultations entre les objets.

La figure 8 protente moins de cas d'interaction, mais illustre la possibilit de rendre rollenissant des objets rolls. De plus, l'animation de la socience permet de

^{2.} on remarquera que l'image de synth Łse contient : objets de synth Łse \times masque + objets reconnus en noir r**@**



FIG. 7 - Image originale et deux images de la source enrichie par un drap de synth£se rouge mobile, un tetraŁdre bleu et une sphŁre verte.

juger de la qualit@u positionnement du cube rose, dont le d@lacement est parfaitement synchrone avec celui de l'objet r@l (la mire) sur lequel il est 'pos@(l'animation est visiblel'àdresse WWW mentionn@en ent@ de l'article).



FIG. 8 - Images de la s φ uence enrichie par un une sphŁre verte mobile et un cube rose statique.

6 Conclusion

Nous avons døtontrøla faisabilitøle l'automatisation de l'insertion d'objets 3D virtuels dans des sø quences vidøl en s'appuyant sur une coopøtion de techniques d'analyse et de synth se d'images.

De plus, cette raisation permet, du point de vue de la recherche en analyse d'images, de valider les algorithmes mis envre, et de poser de nouveaux problŁmes.

Diverses directions de recherche sont ouvertes pour rehâr ces contraintes et aboutiruà systŁme plus performant et plus facile d'utilisation :

• Le calcul en diff@les paramLtres du mouvement d'un objet relativementlà cam@l permettrait de traiter des images non acquises par un systLme entiLrement calibr@u qui contient des objets en mouvement. Le fait de reliêr les contraintes de planarit
 de connaissances sur la sc
 Annotation d'autres types de connaissances sur la sc

3D est un autre axe de recherches futures important. En effet, nous avons opt@our une approche r@ion car la RES n@essite l'obtention d'un modLlelaefois g@i m@ique et photom@ique, relativement complet. Mais, le choix de facettes n'est pas exclusif, et l'association de r@ions avec des approches compl@ientairesbase par exemple de points, segments, courbes et cartes de profondeur, permettrait d'am@orer encore les r@iltats du processus d'analyse. En particulier, on gagnerait en efficacit@e recalage en disposant de plusieurs types de primitives 3D, et la classe des objets que l'on pourrait traiter serait grandement @ndue.

Du point de vue des applications de synthLse d'images, le mange de sources de synthLse et de so quences rolles permet de mettre en valeur le rollisme de la modasation de l'animation et du rendu.

D'autre part, la faisabilit@'un tel m@ange montre qu'il peut s'av@r trLs int@ssant pour les applications pratiques d'exploiter au maximum des scLnes et des @ants r@s, pour n'utiliser des objets virtuels que dà u ils sont vraiment indispensables (ce qu'avait d@ commend@ustrer l'@uipe d'A. Fellous [6]).

Rences

- Feiner (S.), MacIntyre (B.) et Seligmann (D.). Knowledge based augmented reality. *Communications of the ACM*, vol. 36, juillet 1993.
- [2] Caudel (T.) et Mizell (D.W.). Augmented reality: An application of heads-up display technology to manual manufacturing processes. *In* : *Hawaii International Conference on System Sciences*.
- [3] Ahlers (K. H.), Kramer (A.), Breen (D. E.), Chevalier (P. Y.), Crampton (C.), Rose (E.), Tuceryan (M.), Whitaker (R. T.) et Greer (D.). – Distributed augmented reality for collaborative design applications. *In*: *Eurographics'95*, pp. 3–14.
- [4] Whitaker (R. T.), Crampton (C.), Breen (D. E.), Tuceryan (M.) et Rose (E.). – Object calibration for augmented reality. *In: Eurographics'95*, pp. 15–27.
- [5] Bajura (M.), Fuchs (H.) et Ohbuchi (R.). Merging virtual objects with the real world: seeing ultrasound imagery within the patient. *Computer Graphics (SIGGRAPH '92 Proceedings)*, juillet 1992, pp. 203–210.
- [6] Fellous (A.). The STV-synthetic TV system. In: Journees INRIA Analyse-SynthLse d'images, pp. 28–35. – ENST, Paris, janvier 1994.
- [7] Fournier (A.), Gunawan (A. S.) et Romanzin (C.).
 Common illumination between real and computer generated scenes. *In: Proceedings of Graphics Interface '93*. Canadian Information Processing Society, pp. 254–262. Toronto, Ontario, Canada, mai 1993.
- [8] Fournier (A.). Illumination problems in computer augmented reality. *In: Journees INRIA*

Analyse-SynthLse d'images, pp. 1–21. – ENST, Paris, janvier 1994.

- [9] Horaud (Radu) et Dornaika (Fadi). Hand-eye calibration. Int. Journal of Robotics Research, vol. 14, Spring 1995. – A patná.
- [10] Tarel (J.P.) et Gagalowicz (A.). Calibration de cam@ab@ase d'ellipses. *Traitement du Signal*, vol. 12, n 2, 1995, pp. 177–187. – http://wwwrocq.inria.fr/syntim/textes/calib-eng.html.
- Boufama (B.), Mohr (R.) et Veillon (F.). Euclidean constraints for uncalibrated reconstruction.
 In: IEEE International Conference on Computer Vision, pp. 466–470. Berlin, March 1993.
- [12] Tarel (J.P.). Une m@iode de calibration radiom@ique de cam@i@cale variable. In: 10Lme congrLs AFCET, Reconnaissance des Formes et Intelligence Artificielle. – Rennes, France, 1996.
- [14] Sander (Peter T.), Vinet (Laurent), Cohen (Laurent) et Gagalowicz (Andr) Hierarchical regions based stereo matching. *In*: *Proceedings of the Sixth Scandinavian Conference on Image Analysis*, pp. 71–78. Oulu, Finland, June 1989.
- [15] Randriamasy (S.). Segmentation descendante coop@utive en r@ions de paires d'images st@pscopiques. – ThLse de PhD, Universit@paris-IX Dauphine, 1992.
- [16] Randriamasy (S.) et Gagalowicz (A.). Region based stereo matching oriented image processing.
 In: Proceedings of Computer Vision and Pattern Recognition. – Maui, Hawaii, June 1991.
- [17] Boujemaa (Nozha), Stamon (Georges) et Gagalowicz (AndrØ – ModØsation floue pour la segmentation d'images. In: 9Łme congrŁs AFCET, Reconnaissance des Formes et Intelligence Artificielle.
- [18] Ackah-Miezan (A.) et Gagalowicz (A.). Discrete models for energy minimizing segmentation. In: Proceedings of the 3rd International Conference on Computer Vision, pp. 200–207. 11-13 May, Berlin, 1993. http://www-rocq.inria.fr/syntim/textes/iccv93-eng.html.
- [19] Voen (J.M.). Techniques de reconstruction globale par analyse de paires d'images stor scopiques. – ThŁse de PhD, Universit@aris-VII, 1995.
- [20] Tarel (Jean-Philippe) et VØen (Jean-Marc). A generic approach for planar patches stereo reconstruction. In: Proceedings of the Scandinavian Conference on Image Analysis, pp. 1061– 1070. – Uppsala, Sweden, 1995. http://wwwrocq.inria.fr/syntim/textes/scia95-eng.html.

- [21] Koivunen (V.), Vøen (J-M.) et Bajcsy (R.). Procedural cad-models from range data. In: Proceedings of Towards World Class Manufacturing '93.
- [22] Zhang (Z.). Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, vol. 13, n 2, 1994, pp. 119–152. also INRIA Tech. Report #1658.
- [23] Tarel (J.P.) et Boujemaa (N.). Une approche floue du recalage 3d: gaacit@et robustesse. In: 10Lme congrLs AFCET, Reconnaissance des Formes et Intelligence Artificielle. – Rennes, France, 1996.
- [24] Serfaty (V.), Ackah-Miezan (A.), Lutton (E.) et Gagalowicz (A.). – Photometric analysis as an aid to 3D reconstruction of indoor scenes. *In: Image Modeling, L.A. Ray, J.R. Sullivan*, pp. 196–207. – Editors, proc. SPIE 1904.
- [25] Foley (J. D.), van Dam (A.), Feiner (S. K.) et Hughes (J. F.). – Computer Graphics: Principles and Practices (2nd Edition). – Addison Wesley, 1990.
- [26] Watt (A.) et Watt (M.). Advanced Rendering and Animation Techniques: Theory and Practice.
 – Addison-Wesley, 1991.
- [27] Meas-Yedid (V.), Tarel (J.P.) et Gagalowicz (A.).
 Calibration monopole faible et construction interactive de modèles 3D de scènes. *In: Congrès Reconnaissance des Formes et Intelligence Artificielle, (RFIA'94).* AFCET. Paris, France, 1994.
- [28] Provot (X.). Deformation constraints in a massspring model to describe rigid cloth behavior. *In : Proceedings of Graphic Interface*'95, pp. 147– 154. – Qubec, Canada, mai 1995.